

STIC-Biotech/ChemLib

120537

From: Li, Bao-Qun  
Sent: Wednesday, April 28, 2004 7:54 AM  
To: STIC-Biotech/ChemLib

Please do the sequence homology search for SEQ ID NO: 2 of application SN. 10,089,292. Thanks

Bao Qun Li

Art unit 1648, REM, 3D24.

Tel. 20904

RECEIVED  
APR 28 2004  
STIC

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 4/29  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2004, 14:17:48 ; Search time 20 Seconds  
(without alignments)  
1024.440 Million cell updates/sec

Title: US-10-089-292A-2

Perfect score: 1095  
Sequence: 1 QLFYSRPVVSANGPTVKLY.....SLGAGPVSISAVAVLAPPPR 213

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1076	98.3	660	1	VHWH2	structural protein
2	1043	95.3	659	1	B44212	structural protein
3	91.5	8.4	1305	2	AB0168	probable cell divi
4	91.5	8.4	2130	2	AB0821	probable exported
5	90.5	8.3	261	2	S76484	hypothetical prote
6	90	8.2	1045	2	AH3199	endoglucanase B (E
7	89	8.1	291	2	C70143	biphenyl-2,3-diol
8	89	8.1	477	2	C70143	pyruvate kinase (p
9	89	8.1	1802	2	S69703	HKR1 protein precu
10	88.5	8.1	665	2	F95053	cell wall surface
11	88	8.0	583	2	S67571	hypothetical prote
12	87	7.9	329	2	F87318	transcription regu
13	85.5	7.8	817	2	T01866	hypothetical prote
14	85.5	7.8	899	2	T16204	hypothetical prote
15	85	7.8	292	2	A11937	hypothetical prote
16	84.5	7.7	432	2	A1EUC5	adenylosuccinate s
17	84.5	7.7	432	2	A98273	adenylosuccinate s
18	84.5	7.7	432	2	A86114	adenylosuccinate s
19	84.5	7.7	432	2	AF1049	adenylosuccinate s
20	83.5	7.6	253	2	C75611	transcription regu
21	83	7.6	288	2	T21732	hypothetical prote
22	83	7.6	436	2	T03702	hypothetical prote
23	83	7.6	569	2	C91195	hypothetical prote
24	83	7.6	569	2	D86042	hypothetical prote
25	83	7.6	617	2	A56051	mvocyte nuclear fa
26	83	7.6	765	2	C64981	beta-glucosidase (
27	83	7.6	765	2	C91006	beta-D-glucoside g
28	83	7.6	765	2	D85850	beta-D-glucoside g
29	83	7.6	2422	2	T12687	ALR protein homolo

30 82.5 7.5 234 2 S22134 capsid protein - r  
31 82.5 7.5 278 2 G85887 ethanalamine utili  
32 82.5 7.5 278 2 E65020 ethanalamine utili  
33 82.5 7.5 278 2 D91043 ethanalamine utili  
34 82.5 7.5 386 2 S52718 coat protein - car  
35 82 7.5 335 1 G2MSAB Ig gamma-2a chain  
36 82 7.5 1946 2 AE1449 hypothetical prote  
37 81.5 7.4 1077 2 H96007 probable acriflavi  
38 81 7.4 394 2 A70504 translatable ppe prote  
39 81 7.4 418 2 S09153 translation elonga  
40 81 7.4 563 1 CZCLBM cellulase (EC 3.2.  
41 81 7.4 693 2 D90441 ABC transporter, p  
42 81 7.4 868 2 E89897 conserved hypothet  
43 81 7.4 1335 2 T18289 rAGAP protein - s  
44 80.5 7.4 470 2 AF1972 hypothetical prote  
45 80.5 7.4 558 2 T30418 chitinase (EC 3.2.

ALIGNMENTS

RESULT 1

VHWH2

structural protein 2 precursor - hepatitis E virus (strain Burma)

C:Species: hepatitis E virus

C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 23-Jul-1999

C:Accession: C40728

R:Tam, A.W.; Smith, M.M.; Guerra, M.E.; Huang, C.C.; Bradley, D.W.; Fry, K.E.; Reyes, G.F.

Virology 185, 120-131, 1991

A:Title: Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length vi

A:Reference number: A40778; MUID:92024067; PMID:1926770

A:Accession: C40778

A:Molecule type: genomic RNA

A:Residues: 1-660 <TAM>

A:Cross-references: CB:W73218; NID:G330023; PIDN:AAA5736.1; PID:G330026

A:Note: the authors translated the codon:CGC for residue 2 as Ala

C:Superfamily: hepatitis E virus structural protein 2

C:Keywords: structural protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-660/Product: structural protein 2 #status predicted <SP2>

Query Match 98.3%; Score 1076; DB 1; Length 660;

Best Local Similarity 100.0%; Pred. No. 5.1e-85;

Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGPTVKLYTSVENAQDKGIAIPHDIDLGESRVVIQDYNQHEQDRPT 60

Db 394 QLFYSRPVVSANGPTVKLYTSVENAQDKGIAIPHDIDLGESRVVIQDYNQHEQDRPT 453

QY 61 PSPAPSRPFVLRANDVLMSLTAABYDQSTYSGSTGPFVYVSDSVTLNVNATGAQAVARS 120

Db 454 PSPAPSRPFVLRANDVLMSLTAABYDQSTYSGSTGPFVYVSDSVTLNVNATGAQAVARS 513

QY 121 LDWTKVLDGRPLSTIQOYKTFVLPRLGKLSFWEAGTTKAGYPYNNYNTTASDQLLVEN 180

Db 514 LDWTKVLDGRPLSTIQOYKTFVLPRLGKLSFWEAGTTKAGYPYNNYNTTASDQLLVEN 573

QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210

Db 574 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 603

RESULT 2

B44212

structural protein 2 precursor - hepatitis E virus (strain Mexico)

C:Species: hepatitis E virus

C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999

C:Accession: B44212

R:Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.; Bradley, D.W.; Tam, A.W.;

Virology 191, 550-558, 1992

A:Title: Molecular-cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV

A:Reference number: A44212; MUID:93079857; PMID:1448913

A:Accession: B44212

A:Molecule type: genomic RNA

A:Residues: 1-659 <HUA>

A:Cross-references: GB:M74506; NID:g330017; PIDN:AAA5732.1; PID:g330020

C:Superfamily: hepatitis E virus structural protein 2

F:1-22/Domain: signal sequence

F:23-659/Product: structural protein 2 #status predicted <SP2>

Query Match 95.3%; Score 1043; DB 1; Length 659;  
Best Local Similarity 94.3%; Pred. No. 3.6e-82;  
Matches 198; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 QLFYSRPVWSGANGPTVKLYTSVENAQDQKGIAPHDIDLGSRVVIQDYDQHEQDRPT 60  
DB 393 QLFYSRPVWSGANGPTVKLYTSVENAQDQKGIAPHDIDLGSRVVIQDYDQHEQDRPT 452

QY 61 PSPAPSRPFVLRANDVLMVLSLTAAYDQSTGSGTGPVYVSDSVTLVNVATGAQAVARS 120

DB 453 PSPAPSRPFVLRANDVLMVLSLTAAYDQSTGSGTGPVYVSDSVTLVNVATGAQAVARS 512

QY 121 LDWTKVTLGDRPLSTIQOYKTFVLRGLKSFWEAGTTKAGYPYNYNTTASDQLLVEN 180

DB 513 LDWTKVTLGDRPLSTIQOYKTFVLRGLKSFWEAGTTKAGYPYNYNTTASDQLLVEN 572

QY 181 AAGHRVAISTVTTSLGAGPVVISAVAVLAP 210

DB 573 AAGHRVAISTVTTSLGAGPVVISAVAVLAP 602

RESULT 3

AB0168

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001

C:Accession: AB0168

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Molecule type: DNA

A:Residues: 1-1305 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90205.1; PID:g15979425; GSPDB:GN00175

C:Genetics:

A:Gene: ftsK

Query Match 8.4%; Score 91.5; DB 2; Length 1305;

Best Local Similarity 27.9%; Pred. No. 15;

Matches 38; Conservative 17; Mismatches 40; Indels 41; Gaps 7;

QY 14 EPTVKLYTSVENAQDQKGIAPHDIDLGSRVVIQDYDQHEQDRPTFPSP----- 63

DB 664 EPTVKLYTSVENAQDQKGIAPHDIDLGSRVVIQDYDQHEQDRPTFPSP----- 714

QY 64 APSRPVLRANDVLMVLSLTAAYDQSTGSGT-----GPVYVSDSVTLVNVATGAQAVAR 119

DB 715 APSRPVLRANDVLMVLSLTAAYDQSTGSGT-----GPVYVSDSVTLVNVATGAQAVAR 757

QY 120 SLDWTKV-TLDGRPLS 134

DB 758 SVTPTSTASLNTAPVS 773

RESULT 4

AB0821

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C:Accession: AB0821

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulle, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AB0821

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2130 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD02718.1; PID:g16503731; GSPDB:GN00176

C:Genetics:

A:Gene: STV2760

Query Match 8.4%; Score 91.5; DB 2; Length 2130;

Best Local Similarity 24.4%; Pred. No. 28;

Matches 53; Conservative 26; Mismatches 93; Indels 45; Gaps 10;

QY 8 VVSANGPTVKLYTSVENAQDQKGIAPHDIDLGSRVVIQDYDQHEQDRPT 60

DB 1231 IKTAQGWPIIRSYWSGTNAGTTPGAPPDYDYTVLDNDAAHRKVPNISDSRQVQICAAAT 1290

QY 61 PSPAPSRPFVLRANDVLMVLSLTAAYDQSTGSGTGPVYVSDSVTLVNVATGAQAVARS 120

DB 1291 PQPLAGR-----ITLTSTLATSDIOAVKAKN-----SDSIPL--VITTTDAAGNP 1334

QY 121 LDWTKVTLGDRPLSTIQOYKTFVLRGLKSFWEAGTTKAGYPYNYN-----TTASD 174

DB 1335 VPYTPFSLI-RDAGTARNTSYTF-----TGSTNNMLAPPTGSAQQFYNGTYTGATGAD 1388

QY 175 QLLVANAAGHRVAISTVTTSLGAGPVVISAVAVLAP 211

DB 1389 -----GTAVLTITQAAGPVKNVITALTDP 1415

RESULT 5

S76484

C:Species: Synecocystis sp.

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76484

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-261 <KAN>

A:Cross-references: EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BA18613.1; PID:d1019346

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Start codon: GTG

Query Match 8.3%; Score 90.5; DB 2; Length 261;

Best Local Similarity 23.6%; Pred. No. 2.2; Indels 69; Gaps 14;

Matches 54; Conservative 32; Mismatches 74; Indels 69; Gaps 14;

QY 2 LFYSRPVWSGANGPTVKLYTSVENAQDQKGIAPHDIDLGSRVVIQDYDQHEQDRPT 54

DB 20 LLYVSLPAIANNNGSKQALPELQIHP-----LPELVLDLGISLENRSLLGD-DGTD 67

QY 55 EQDRPTSPAPSRPFVLRANDVLMVLSL-----TAARYDQS-TYGSSTGPVYVSDSVTLVNV 110

DB 68 YFSQVQSP-----LGLYLSKFPVTVAVDPPGLTPGSAQAQRY----- 108

QY 111 ATGAQAVARS-LDWTKVTLGDRPLSTIQOYKSK-----TFF-----VLPURGLKLSFEWA 157

DB 109 -TWOQAIAKTAAIDMQEFFF-----PLTIVTEADITIFYREPPPLPRIVDPTGLVSGRA 163



[illegible]

RESULT 14  
T16204  
hypothetical protein F28F5.3 - *Caenorhabditis elegans*  
C Species: *Caenorhabditis elegans*  
C Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 04-Mar-2000  
C Accession: T16204  
R; Pauley, A.  
submitted to the EMBL Data Library, April 1994  
A Description: The sequence of C. elegans cosmid F28F5.  
A Reference number: Z18477  
A Accession: T16204  
A Status: preliminary; translated from GB/EMBL/DBJ

Search completed: April 28, 2004, 14:21:53  
Job time : 21 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2004, 14:17:13 ; Search time 45 Seconds  
(without alignments)  
1493.453 Million cell updates/sec

Title: US-10-089-292A-2  
Perfect score: 1095  
Sequence: 1 QLFYSRPVVSANGEPTVKLY.....SLGAGPVSISAVLAPPPR 213

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL 25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mic.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1076	98.3	344	12 Q9WJZ5	Q9WJZ5 hepatitis e
2	1076	98.3	605	12 Q9VWL0	Q9VWL0 hepatitis e
3	1076	98.3	660	12 Q9ESG5	Q9ESG5 hepatitis e
4	1076	98.3	660	12 Q9W7W9	Q9W7W9 hepatitis e
5	1072	97.9	659	12 Q8JT32	Q8JT32 hepatitis e
6	1072	97.9	660	12 Q69411	Q69411 hepatitis e
7	1072	97.9	660	12 Q89468	Q89468 hepatitis e
8	1072	97.9	660	12 Q9WQAO	Q9WQAO hepatitis e
9	1072	97.9	660	12 Q91855	Q91855 hepatitis e
10	1071	97.8	660	12 Q91856	Q91856 hepatitis e
11	1071	97.8	660	12 Q81871	Q81871 hepatitis e
12	1071	97.8	660	12 Q81878	Q81878 hepatitis e
13	1069	97.6	660	12 Q9WLL4	Q9WLL4 hepatitis e
14	1067	97.4	525	12 Q39947	Q39947 hepatitis e
15	1066	97.4	660	12 Q69419	Q69419 hepatitis e
16	1065	97.3	436	12 Q9WI48	Q9WI48 hepatitis e

17	1051	96.0	660	12 Q68985	Q68985 hepatitis e
18	1036	94.6	660	12 Q801Y5	Q801Y5 hepatitis e
19	1033	94.3	660	12 Q9YLQ9	Q9YLQ9 hepatitis e
20	1033	94.3	660	12 Q36613	Q36613 swine hepat
21	1033	94.3	660	12 Q8JQM4	Q8JQM4 hepatitis e
22	1033	94.3	660	12 Q8JQM8	Q8JQM8 hepatitis e
23	1029	94.0	660	12 Q91114	Q91114 hepatitis e
24	1029	94.0	660	12 Q8V729	Q8V729 swine hepat
25	1029	94.0	660	12 Q8AZM3	Q8AZM3 hepatitis e
26	1029	94.0	660	12 Q801R7	Q801R7 hepatitis e
27	1028	93.9	258	12 Q81861	Q81861 hepatitis e
28	1026	93.7	550	12 Q8JVJ3	Q8JVJ3 hepatitis e
29	1026	93.7	671	12 Q8JQM1	Q8JQM1 hepatitis e
30	1026	93.7	674	12 Q8JUN6	Q8JUN6 hepatitis e
31	1024	93.5	674	12 Q8JJN2	Q8JJN2 hepatitis e
32	1024	93.5	674	12 Q7TGF1	Q7TGF1 hepatitis e
33	1023	93.4	486	12 Q9WNN2	Q9WNN2 hepatitis e
34	1023	93.4	674	12 Q805E0	Q805E0 swine hepat
35	1023	93.4	674	12 Q806D7	Q806D7 hepatitis e
36	1023	93.4	674	12 Q801R5	Q801R5 hepatitis e
37	1022	93.3	660	12 Q8BB05	Q8BB05 swine hepat
38	1020	93.2	660	12 Q9YLR2	Q9YLR2 hepatitis e
39	1019	93.1	466	12 Q9WNN1	Q9WNN1 hepatitis e
40	1011	92.3	672	12 Q9IVZ8	Q9IVZ8 hepatitis e
41	1010	92.2	248	12 Q81860	Q81860 hepatitis e
42	630	57.5	184	12 Q9YPB0	Q9YPB0 hepatitis e
43	512	46.8	149	12 Q9W9E8	Q9W9E8 hepatitis e
44	512	46.8	149	12 Q71348	Q71348 hepatitis e
45	505	46.1	149	12 Q56047	Q56047 hepatitis e

## ALIGNMENTS

## RESULT 1

Q9WJZ5 PRELIMINARY; PRT; 344 AA.

AC Q9WJZ5;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Structural protein 2 (Fragment).  
OS Hepatitis E virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage;  
OC Hepatitis E-like viruses.  
OX NCBI\_TaxID=12461;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2B Abbottabad;  
RA Van Cuyck-Gandre H., Clements N.J., Zhang H.Y., Caudill J.C.,  
RA Cohen S.G., Coursaget P., Buissson Y., Warren R.L., Longer C.F.;  
RT "Partial Sequence of HEV Isolates from North Africa and Pakistan:  
RT Comparison with Known HEV Sequences."  
RRL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
VDR EMBL; U40044; AAD09445.1; -.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR004261; SP2.  
DR Pfam; PF03014; SP2; 1.  
FT NON TER  
SQ SEQUENCE 344 AA; 37146 MW; D7F3D2D9BB44098B CRC64;

Query Match 98.3%; Score 1076; DB 12; Length 344;  
Best Local Similarity 100.0%; Pred. No. 1e-89;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLFYSRPVVSANGEPTVKLYTSVENAQDQKGIATPHIDIGESRVVTDYDQHQDRPT 60  
Db 78 QLFYSRPVVSANGEPTVKLYTSVENAQDQKGIATPHIDIGESRVVTDYDQHQDRPT 137  
Qy 61 PSPAPSPFVLRANDVWLWLSLTAAYDQSTGYSGSTGPTVYVSDSVTLNVNATGAQAARS 120  
Db 138 PSPAPSPFVLRANDVWLWLSLTAAYDQSTGYSGSTGPTVYVSDSVTLNVNATGAQAARS 197

Qy 121 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
Db 198 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 257  
Qy 181 AAGRVAISTYTTSLGAGPVSISAVAVLAP 210  
Db 258 AAGRVAISTYTTSLGAGPVSISAVAVLAP 287

RESULT 2  
Q9YWL0 PRELIMINARY; PRT; 605 AA.  
ID Q9YWL0  
AC Q9YWL0;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Structural protein (fragment).  
OS Hepatitis E virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage;  
OC Hepatitis E-like viruses.  
OX NCBI\_TaxID=12461;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCN;  
RX MEDLINE=99013657; PubMed=9797311;  
RA Pina S., Jofre J., Emerson S.U., Purcell R.H., Girones R.;  
RT "Characterization of a strain of infectious hepatitis E virus isolated  
from sewage in an area where hepatitis E is not endemic.";  
RL Appl. Environ. Microbiol. 64:4485-4488 (1998).  
DR EMBL; AF058684; AAC77808.1; -;  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR004261; SP2.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF03014; SP2; 1.  
FT NON\_TER 605 605  
SQ SEQUENCE 605 AA; 64740 MW; A088F76F874B2B72 CRC64;

Query Match 98.3%; Score 1076; DB 12; Length 605;  
Best Local Similarity 100.0%; Pred. No. 2.4e-89;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGSRVVIQDYDNQHEQDRPT 60  
Db 394 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGSRVVIQDYDNQHEQDRPT 453

Qy 61 PSPAPSPFSVLRANDVLMLSLTAAYDQSTYGSSTGPVYVSDSVTLVNVATGAQAVARS 120  
Db 454 PSPAPSPFSVLRANDVLMLSLTAAYDQSTYGSSTGPVYVSDSVTLVNVATGAQAVARS 513

Qy 121 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
Db 514 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 573

Qy 181 AAGRVAISTYTTSLGAGPVSISAVAVLAP 210  
Db 574 AAGRVAISTYTTSLGAGPVSISAVAVLAP 603

RESULT 3  
Q9E8G5 PRELIMINARY; PRT; 660 AA.  
ID Q9E8G5  
AC Q9E8G5;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Structural protein.  
OS Hepatitis E virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage;  
OC Hepatitis E-like viruses.  
OX NCBI\_TaxID=12461;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Abb-2B;

RX MEDLINE=20271579; PubMed=10813471;  
RA van Cuyck-Gandre H., Zhang H.Y., Tsarev S.A., Warren R.L.,  
RA Caudill J.D., Snellings N.J., Begot L., Innis B.L., Longer C.F.;  
RT "Phylogenetically distinct hepatitis E viruses in Pakistan.";  
RL Am. J. Trop. Med. Hyg. 62:187-189 (2000).  
DR EMBL; AF185822; AAG16766.1; -;  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR004261; SP2.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF03014; SP2; 1.  
SQ SEQUENCE 660 AA; 70903 MW; 1F506BE3CFB3BACE CRC64;

Query Match 98.3%; Score 1076; DB 12; Length 660;  
Best Local Similarity 100.0%; Pred. No. 2.7e-89;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGSRVVIQDYDNQHEQDRPT 60  
Db 394 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGSRVVIQDYDNQHEQDRPT 453

Qy 61 PSPAPSPFSVLRANDVLMLSLTAAYDQSTYGSSTGPVYVSDSVTLVNVATGAQAVARS 120  
Db 454 PSPAPSPFSVLRANDVLMLSLTAAYDQSTYGSSTGPVYVSDSVTLVNVATGAQAVARS 513

Qy 121 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
Db 514 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 573

Qy 181 AAGRVAISTYTTSLGAGPVSISAVAVLAP 210  
Db 574 AAGRVAISTYTTSLGAGPVSISAVAVLAP 603

## RESULT 4

Q9W7W9 PRELIMINARY; PRT; 660 AA.  
ID Q9W7W9  
AC Q9W7W9;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Structural protein 2 (Structural viral protein).  
OS Hepatitis E virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage;  
OC Hepatitis E-like viruses.  
OX NCBI\_TaxID=12461;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HEV-Morocco;  
RX MEDLINE=99107414; PubMed=9892396;  
RA Meng J., Cong M., Dai X., Pillot J., Purdy M.A., Fields H.A.,  
RA Khudvakov Y.E.;  
RT "Primary structure of open reading frame 2 and 3 of the hepatitis E  
virus isolated from Morocco.";  
RL J. Med. Virol. 57:126-133 (1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Morocco;  
RA Chen G., Meng J.;  
RT "Identification of the 5' Capped and 3' Complete Terminal Sequence of  
the Hepatitis E virus isolated from Morocco.";  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF065061; AAD10627.1; -;  
DR EMBL; AY230202; AAO72992.1; -;  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR004261; SP2.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF03014; SP2; 1.  
SQ SEQUENCE 660 AA; 70943 MW; 9719D30CCD932950 CRC64;

Query Match 98.3%; Score 1076; DB 12; Length 660;  
Best Local Similarity 100.0%; Pred. No. 2.7e-89;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 QLFYSRPVVSANGPTVKLYTSVENAQDQKGIAPHIDILGESRVVIQDYNQHEQDRPT 60
Db 394 QLFYSRPVVSANGPTVKLYTSVENAQDQKGIAPHIDILGESRVVIQDYNQHEQDRPT 453
Qy 61 PSPAPSRPFSVLRANDVWLMLSLTAAEYDQSTYGSSTGTPVYSDSVTLNVNATGAQAVARS 120
Db 454 PSPAPSRPFSVLRANDVWLMLSLTAAEYDQSTYGSSTGTPVYSDSVTLNVNATGAQAVARS 513
Qy 121 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180
Db 514 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 573
Qy 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
Db 574 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 603

RESULT 5
Q8JT32 PRELIMINARY; PRT; 659 AA.
AC 08JT32;
DB 08JT32;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072932; PubMed=12076829;
RA Jameel S., Zafrullah M., Chawla Y.K., Dilawari J.B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF459438; AAME6330.1; -.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR Pfam: PF03014; SP2; 1.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 659 AA; 70842 MW; ABCC177EBE6012C8 CRC64;

Query Match 97.9%; Score 1072; DB 12; Length 659;
Best Local Similarity 99.5%; Pred. No. 6.3e-89;
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QLFYSRPVVSANGPTVKLYTSVENAQDQKGIAPHIDILGESRVVIQDYNQHEQDRPT 60
Db 393 QLFYSRPVVSANGPTVKLYTSVENAQDQKGIAPHIDILGESRVVIQDYNQHEQDRPT 452
Qy 61 PSPAPSRPFSVLRANDVWLMLSLTAAEYDQSTYGSSTGTPVYSDSVTLNVNATGAQAVARS 120
Db 453 PSPAPSRPFSVLRANDVWLMLSLTAAEYDQSTYGSSTGTPVYSDSVTLNVNATGAQAVARS 512
Qy 121 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180
Db 513 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 572
Qy 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
Db 573 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 602

RESULT 6
Q69411 PRELIMINARY; PRT; 660 AA.
AC Q69411
AC Q69411;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=hev037;
RA Donati M.C., Fagan E.A., Harrison T.J.;
RT "Sequence analysis of full length HEV clones derived directly from human liver in fulminant hepatitis E.";
RL (In) Rizzetto M., Purcell R.H., Gerin J.L., Verme G (eds.);
RL VIRAL HEPATITIS AND LIVER DISEASE, pp.313-316, Edizioni Minerva Medica, Torino (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=hev037;
RA Harrison T.J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: X98292; CAA66937.1; -.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71024 MW; F37E3FEFF7A4EAD1 CRC64;

Query Match 97.9%; Score 1072; DB 12; Length 660;
Best Local Similarity 99.0%; Pred. No. 6.3e-89;
Matches 208; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLFYSRPVVSANGPTVKLYTSVENAQDQKGIAPHIDILGESRVVIQDYNQHEQDRPT 60
Db 394 QLFYSRPVVSANGPTVKLYTSVENAQDQKGIAPHIDILGESRVVIQDYNQHEQDRPT 453
Qy 61 PSPAPSRPFSVLRANDVWLMLSLTAAEYDQSTYGSSTGTPVYSDSVTLNVNATGAQAVARS 120
Db 454 PSPAPSRPFSVLRANDVWLMLSLTAAEYDQSTYGSSTGTPVYSDSVTLNVNATGAQAVARS 513
Qy 121 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180
Db 514 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 573
Qy 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
Db 574 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 603

RESULT 7
Q89468 PRELIMINARY; PRT; 660 AA.
AC Q89468
AC Q89468;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE UNNAMED protein product.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K52-87;
RA Yin S.R., Purcell R.H., Emerson S.U.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K52-87;
RX MEDLINE=95176571; PubMed=7871758;
RA Yin S., Purcell R.H., Emerson S.U.;
RT "A new Chinese isolate of hepatitis E virus: comparison with strains recovered from different geographical regions.";
RL Virus Genes 9:23-32 (1994).
```

```
DR EMBL; L25547; AAA91080.1; --
DR EMBL; L25595; AAA65490.1; --
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70979 MW; DC68116DCD639175 CRC64;

Query Match          97.9%; Score 1072; DB 12; Length 660;
Best Local Similarity 99.5%; Pred. No. 6.3e-89;
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHEQDRPT 60
DB 394 QLFYSRPVVSANGEPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHEQDRPT 453
QY 61 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 120
DB 454 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 513
QY 121 LDWTKVTLDGRPLSTIQOYKSTFFVPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180
DB 514 LDWTKVTLDGRPLSTIQOYKSTFFVPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 573
QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
DB 574 AAGHRVAISTYTTSLGAGPVSISAVAVLTP 603

RESULT 8
Q9WQAO PRELIMINARY; PRT; 660 AA.
AC Q9WQAO;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF-2.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKL-90;
RX MEDLINE=95350000; PubMed=10423137;
RA Arankalle V.A., Paranjape S., Emerson S.U., Purcell R.H.,
RA Walimbe A.M.;
RT "Phylogenetic analysis of hepatitis E virus isolates from India (1976-1993).";
RL J. Gen. Virol. 80:1691-1700(1999).
DR EMBL; AF124407; AAD45493.1; --
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71041 MW; 03B72DDF0AB7B521 CRC64;

Query Match          97.9%; Score 1072; DB 12; Length 660;
Best Local Similarity 99.5%; Pred. No. 6.3e-89;
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHEQDRPT 60
DB 394 QLFYSRPVVSANGEPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHEQDRPT 453
QY 61 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 120
DB 454 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 513
QY 121 LDWTKVTLDGRPLSTIQOYKSTFFVPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180
DB 514 LDWTKVTLDGRPLSTIQOYKSTFFVPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 573
```

```
QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
DB 574 AAGHRVAISTYTTSLGAGPVSISAVAVFAP 603

RESULT 9
Q91855 PRELIMINARY; PRT; 660 AA.
AC Q91855;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=93-Egypt;
RX MEDLINE=99105430; PubMed=9890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,
RA van Cuyck-Gandre H., Longer C.F., Innis B.L.;
RT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
RL J. Med. Virol. 57:68-74(1999).
DR EMBL; AF051351; AAC35761.1; --
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71040 MW; 044FF5EA7C492791 CRC64;

Query Match          97.9%; Score 1072; DB 12; Length 660;
Best Local Similarity 99.0%; Pred. No. 6.3e-89;
Matches 208; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHEQDRPT 60
DB 394 QLFYSRPVVSANGEPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHEQDRPT 453
QY 61 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 120
DB 454 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 513
QY 121 LDWTKVTLDGRPLSTIQOYKSTFFVPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180
DB 514 LDWTKVTLDGRPLSTIQOYKSTFFVPLRGKLSFWEAGTTKAGYPYNYNTTSSDQLLIEN 573
QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
DB 574 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 603

RESULT 10
Q91856 PRELIMINARY; PRT; 660 AA.
AC Q91856;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=94-Egypt;
RX MEDLINE=99105430; PubMed=9890424;
RA Tsarev S.A., Binn L.N., Gomatos P.J., Arthur R.R., Monier M.K.,
RA van Cuyck-Gandre H., Longer C.F., Innis B.L.;
RT "Phylogenetic analysis of hepatitis E virus isolates from Egypt.";
RL J. Med. Virol. 57:68-74(1999).
```

```

DR EMBL: AF051352; AAC35764.1;
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 71026 MW; 16C560FA16941F2A CRC64;

Query Match 97.8%; Score 1071; DB 12; Length 660;
Best Local Similarity 99.0%; Pred. No. 7.7e-89;
Matches 208; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QLFYSRPVWSANGPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYNQHEQDRPT 60
Db |||||||
Qy 394 QLFYSRPVWSANGPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYNQHEQDRPT 453
Db |||||||
Qy 61 PSPAPSRPFSVLRANDVLMVLSLTAAYDQSTYGSSTGTPVYVSDSVTLNVNATGAQAVARS 120
Db |||||||
Qy 454 PSPAPSRPFSVLRANDVLMVLSLTAAYDQSTYGSSTGTPVYVSDSVTLNVNATGAQAVARS 513
Db |||||||
Qy 121 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKSFWEAGTTKAGYPYNTTASDQLLVEN 180
Db |||||||
Qy 514 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKSFWEAGTTKAGYPYNTTASDQLLVEN 573
Db |||||||
Qy 181 AAGHRVAISTVTSISLGGAPVSIASAVAVLAP 210
Db |||||||
Qy 574 ATGHRVAISTVTSISLGGAPVSIASAVAVLAP 603
Db |||||||

RESULT 11
Q81871 ID Q81871 PRELIMINARY; PRT; 660 AA.
AC Q81871;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF 2 precursor.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024067; PubMed=1926770;
RA Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
RA Fry K.E., Reyes G.R.;
RT "Hepatitis E virus (HEV): molecular cloning and sequencing of the
RT full-length viral genome.";
RL Virology 185:120-131(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92261377; PubMed=1584074;
RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,
RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K., Win K.M.;
RA "Hepatitis E virus: cDNA cloning and expression.";
RL Microbiol. Immunol. 36:67-79(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92335008; PubMed=1630924;
RA Aye T.T., Uchida T., Ma X.Z., Iida F., Shikata T., Zhuang H.,
RA Win K.M.;
RT "Complete nucleotide sequence of a hepatitis E virus isolated from the
RT Xinjiang epidemic (1986-1988) of China.";
RL Nucleic Acids Res. 20:3512-3512(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92115700; PubMed=1731327;
RA Tsarev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,
RA Malik I.A., Iqbal M., Purcell R.H.;
RT "Characterization of a prototype strain of hepatitis E virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=92271462; PubMed=1589964;

```

```

RA Fry K.E., Tam A.W., Smith M.W., Kim J.P., Luk K.-C., Young L.M.,
RA Piatak M., Feldman R.A., Yun K.Y., Purdy M.A., McCaustland K.A.,
RA Bradley D.W., Reyes G.R.;
RT "Hepatitis E virus (HEV): Strain variation in the nonstructural gene
RT region encoding consensus motifs for an RNA-dependent RNA polymerase
RT and an ATP/GTP binding site.";
RL Virus Genes 6:173-185(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93348763; PubMed=8346669;
RA Bi S.-L., Purdy M.A., McCaustland K.A., Margolis H.S., Bradley D.W.;
RT "The sequence of hepatitis E virus isolated directly from a single
RT source during an outbreak in China.";
RL Virus Res. 28:233-247(1993).
DR EMBL: L08816; AAA03191.1; -.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
KW Signal.
FT SIGNAL
SQ SEQUENCE 660 AA; 70979 MW; B58F23955FDD6614 CRC64;

Query Match 97.8%; Score 1071; DB 12; Length 660;
Best Local Similarity 99.5%; Pred. No. 7.7e-89;
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QLFYSRPVWSANGPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYNQHEQDRPT 60
Db |||||||
Qy 394 QLFYSRPVWSANGPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYNQHEQDRPT 453
Db |||||||
Qy 61 PSPAPSRPFSVLRANDVLMVLSLTAAYDQSTYGSSTGTPVYVSDSVTLNVNATGAQAVARS 120
Db |||||||
Qy 454 PSPAPSRPFSVLRANDVLMVLSLTAAYDQSTYGSSTGTPVYVSDSVTLNVNATGAQAVARS 513
Db |||||||
Qy 121 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKSFWEAGTTKAGYPYNTTASDQLLVEN 180
Db |||||||
Qy 514 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKSFWEAGTTKAGYPYNTTASDQLLVEN 573
Db |||||||
Qy 181 AAGHRVAISTVTSISLGGAPVSIASAVAVLAP 210
Db |||||||
Qy 574 AAGHRVAISTVTSISLGGAPVSIASAVAVLAP 603
Db |||||||

RESULT 12
Q81878 ID Q81878 PRELIMINARY; PRT; 660 AA.
AC Q81878;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Complete genome sequence.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Uigh179;
RA Uchida T.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: D11093; BAA01867.1; -.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR004261; SP2.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70937 MW; 5412313F1A2EF4A9 CRC64;

Query Match 97.8%; Score 1071; DB 12; Length 660;
Best Local Similarity 99.5%; Pred. No. 7.7e-89;
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QLFYSRPVWSANGPTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYNQHEQDRPT 60

```

```
|||||
394 QLFYSRRVWSANGPTVKLYTSVENAQODKGIAIPHDIDLGESRWIQDYDNQHEQDRPT 453
|||||
61 PSPAPSPFVSLRANVLMWLSLTAAYDQSTYSGSTGPGVYVSDSVTLVNVATGAQAVARS 120
|||||
454 PSPAPSPFVSLRANVLMWLSLTAAYDQSTYSGSTGPGVYVSDSVTLVNVATGAQAVARS 513
|||||
121 LDWTKVTLDRPLSTIIQQYKTFVPLRGLKSPWEAGTTKAGYPYNYNTTASDQLAVEN 180
|||||
514 LDWTKVTLDRPLSTIIQQYKTFVPLRGLKSPWEAGTTKAGYPYNYNTTASDQLAVEN 573
|||||
181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
|||||
574 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 603
|||||

RESULT 13
Q9WLL4 ID Q9WLL4 PRELIMINARY; PRT; 660 AA.
AC Q9WLL4
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK15/92;
RX MEDLINE=99049628; PubMed=9833882;
RA Gouvea V.; Snellings N.; Popek M.J.; Longer C.F.; Innis B.L.;
RT "Hepatitis E virus: complete genome sequence and phylogenetic analysis
of a Nepali isolate.";
RL Virus Res. 57:21-26(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TK15/92;
RA Gouvea V.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF051830; AAC97188.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70798 MW; A04C0185ACC085DB CRC64;

Query Match 97.6%; Score 1069; DB 12; Length 660;
Best Local Similarity 99.5%; Pred. No. 1.2e-88;
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLFYSRRVWSANGPTVKLYTSVENAQODKGIAIPHDIDLGESRWIQDYDNQHEQDRPT 60
|||||
394 QLFYSRRVWSANGPTVKLYTSVENAQODKGIAIPHDIDLGESRWIQDYDNQHEQDRPT 453
|||||
61 PSPAPSPFVSLRANVLMWLSLTAAYDQSTYSGSTGPGVYVSDSVTLVNVATGAQAVARS 120
|||||
454 PSPAPSPFVSLRANVLMWLSLTAAYDQSTYSGSTGPGVYVSDSVTLVNVATGAQAVARS 513
|||||
121 LDWTKVTLDRPLSTIIQQYKTFVPLRGLKSPWEAGTTKAGYPYNYNTTASDQLAVEN 180
|||||
514 LDWTKVTLDRPLSTIIQQYKTFVPLRGLKSPWEAGTTKAGYPYNYNTTASDQLAVEN 573
|||||
181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
|||||
574 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 603
|||||

RESULT 14
O39947 ID O39947 PRELIMINARY; PRT; 525 AA.
AC O39947
```

```
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Structural protein 2 (Fragment).
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chad T3;
RA Van Cuyck-Gandre H.; Caudill J.; Clements N.; Zhang H.; Buisson Y.;
RA Cohen S.; Warren R.; Longer C.;
RT "Partial sequence of HEV isolate from Chad outbreak 1983.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U62121; AAB61825.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
FT NON TER 1
SQ SEQUENCE 525 AA; 56594 MW; 9AFEF48C206351C8 CRC64;

Query Match 97.4%; Score 1067; DB 12; Length 525;
Best Local Similarity 98.6%; Pred. No. 1.3e-88;
Matches 207; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QLFYSRRVWSANGPTVKLYTSVENAQODKGIAIPHDIDLGESRWIQDYDNQHEQDRPT 60
|||||
259 QLFYSRRVWSANGPTVKLYTSVENAQODKGIAIPHDIDLGESRWIQDYDNQHEQDRPT 318
|||||
61 PSPAPSPFVSLRANVLMWLSLTAAYDQSTYSGSTGPGVYVSDSVTLVNVATGAQAVARS 120
|||||
319 PSPAPSPFVSLRANVLMWLSLTAAYDQSTYSGSTGPGVYVSDSVTLVNVATGAQAVARS 378
|||||
121 LDWTKVTLDRPLSTIIQQYKTFVPLRGLKSPWEAGTTKAGYPYNYNTTASDQLAVEN 180
|||||
379 LDWTKVTLDRPLSTIIQQYKTFVPLRGLKSPWEAGTTKAGYPYNYNTTASDQLAVEN 438
|||||
181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210
|||||
439 VAGHRVAISTYTTSLGAGPVSISAVAVLAP 468
|||||

RESULT 15
Q69419 ID Q69419 PRELIMINARY; PRT; 660 AA.
AC Q69419
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF1, ORF2 & ORF3.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RA von Brunn A.; Seebach J.; Thyagarajan S.P.; Mohanavalli B.; Menon T.;
RA Froesner G.;
RT "PCR amplification, cloning and sequence determination of a hepatitis
E virus isolate from Madras, India.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X99441; CAA67804.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR004261; SP2.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF03014; SP2; 1.
SQ SEQUENCE 660 AA; 70960 MW; DAC3DF95F91689F8 CRC64;

Query Match 97.4%; Score 1066; DB 12; Length 660;
Best Local Similarity 99.0%; Pred. No. 2.2e-88;
Matches 208; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy	1	QLFYSRPVVSANGPPTVKLYTSVENAQODKGIAIPHDIDLGESRVVIQDYDNQHEQRPT	60
Db	394	QLFYSRPVVSANGPPTVKLYTSVENAQODKGIAIPHDIDLGESRVVIQDYDNQHEQRPT	453
Qy	61	PSPAPSRPFVLRANDVWLWLSLTAAYDQSTYGSSTGPPVYSDSVTLNVNATGAQAVARS	120
Db	454	PSPAPSRPFVLRANDVWLWLSLTAAYDQSTYGSSTGPPVYSDSVTLNVNATGAQAVARS	513
Qy	121	LDWTKVTLDGRPLSTIQQYSKTFVFLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN	180
Db	514	LDWTKVTLDGRPLSTIQQYSKTFVFLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN	573
Qy	181	AAGHRVAISTYTTSLGAGPVSISAVAVLAP	210
Db	574	AAGHRVAISTYTTSLGAGPVSISAVAVLAP	603

Search completed: April 28, 2004, 14:21:21  
Job time : 47 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2004, 14:18:43 Search time 22 Seconds  
(without alignments)  
499.833 Million cell updates/sec

Title: US-10-089-292A-2

Perfect score: 1095

Sequence: 1 QLFYSRPVVSANGEPTVKLY.....SLGAGPVVISAVAVLAPPPR 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1076	98.3	267	4	US-09-172-699-18
2	1076	98.3	327	1	US-08-240-049B-13
3	1076	98.3	327	1	US-08-259-148A-15
4	1076	98.3	327	1	US-08-484-054-15
5	1076	98.3	327	2	US-07-876-941A-15
6	1076	98.3	327	3	US-08-542-634-17
7	1076	98.3	327	3	US-08-477-292-17
8	1076	98.3	327	4	US-07-870-985A-15
9	1076	98.3	327	5	PCT-US95-13703-17
10	1076	98.3	436	1	US-08-259-148A-17
11	1076	98.3	436	1	US-08-484-054-17
12	1076	98.3	436	2	US-07-876-941A-17
13	1076	98.3	436	4	US-07-870-985A-17
14	1076	98.3	525	3	US-08-542-634-27
15	1076	98.3	525	5	PCT-US95-13703-27
16	1076	98.3	540	3	US-08-542-634-25
17	1076	98.3	540	5	PCT-US95-13703-25
18	1076	98.3	549	3	US-08-542-634-15
19	1076	98.3	549	3	US-08-477-292-15
20	1076	98.3	549	5	PCT-US95-13703-15
21	1076	98.3	552	4	US-09-172-699-16
22	1076	98.3	561	4	US-09-172-699-20
23	1076	98.3	660	1	US-08-240-049B-15
24	1076	98.3	660	1	US-08-259-148A-19
25	1076	98.3	660	1	US-08-484-054-19
26	1076	98.3	660	2	US-07-876-941A-19
27	1076	98.3	660	3	US-08-840-316-2

28	1076	98.3	660	3	US-08-478-507-8	Sequence 8, Appli
29	1076	98.3	660	3	US-08-809-523-2	Sequence 2, Appli
30	1076	98.3	660	3	US-08-542-634-13	Sequence 13, Appl
31	1076	98.3	660	3	US-09-128-275A-8	Sequence 8, Appli
32	1076	98.3	660	3	US-08-471-971-2	Sequence 2, Appli
33	1076	98.3	660	3	US-08-477-292-13	Sequence 13, Appl
34	1076	98.3	660	4	US-09-553-427-8	Sequence 8, Appli
35	1076	98.3	660	4	US-09-462-606-12	Sequence 12, Appl
36	1076	98.3	660	4	US-09-462-606-50	Sequence 50, Appl
37	1076	98.3	660	4	US-07-870-985A-19	Sequence 19, Appl
38	1076	98.3	660	4	US-09-402-776-2	Sequence 2, Appli
39	1076	98.3	660	4	US-09-172-699-2	Sequence 2, Appli
40	1076	98.3	660	5	PCT-US93-08849A-2	Sequence 2, Appli
41	1076	98.3	660	5	PCT-US93-08849-2	Sequence 2, Appli
42	1076	98.3	660	5	PCT-US95-13703-13	Sequence 13, Appl
43	1072	97.9	660	4	US-09-462-606-49	Sequence 49, Appl
44	1072	97.9	660	4	US-09-462-606-53	Sequence 53, Appl
45	1071	97.8	660	4	US-09-462-606-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1

US-09-172-699-18  
; Sequence 18, Application US/09172699A  
; Patent No. 6514690  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David A.  
; APPLICANT: Locarnini, Stephen A.  
; APPLICANT: Toressi, Joseph  
; APPLICANT: Hu, Zhuang  
; APPLICANT: Li, Fan  
; TITLE OF INVENTION: IMMUNOREACTIVE ANTIGENS OF HEPATITIS E VIRUS  
; FILE REFERENCE: Davies Col. Cave  
; CURRENT APPLICATION NUMBER: US/09/172,699A  
; CURRENT FILING DATE: 1998-10-14  
; EARLIER APPLICATION NUMBER: 08/617,927  
; EARLIER FILING DATE: 1996-06-20  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 267  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-172-699-18

Query Match	98.3%	Score 1076;	DB 4;	Length 267;
Best Local Similarity	100.0%;	Pred. No. 3.2e-104;		
Matches 210;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDIGESRVVIQDYDNOHQDRPT	60	
Db	1	QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDIGESRVVIQDYDNOHQDRPT	60	
QY	61	PSAPSRPFVSLRANDVWLSTAAEYDQSTYSGSTGPTVVSSTVTLVNVATGAQAVARS	120	
Db	61	PSAPSRPFVSLRANDVWLSTAAEYDQSTYSGSTGPTVVSSTVTLVNVATGAQAVARS	120	
QY	121	LDWTKVTLDRPLSTTQQYSKTFVPLRKLKSFWEAGTTKAGYPYNYNTTASDQLLVEN	180	
Db	121	LDWTKVTLDRPLSTTQQYSKTFVPLRKLKSFWEAGTTKAGYPYNYNTTASDQLLVEN	180	
QY	181	AAGHRAVISTYTTSLGAGPVVISAVAVLAP	210	
Db	181	AAGHRAVISTYTTSLGAGPVVISAVAVLAP	210	

RESULT 2

US-08-240-049B-13  
; Sequence 13, Application US/08240049B  
; Patent No. 5686239  
; GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R.  
APPLICANT: Tam, Albert W.  
APPLICANT: Yarbough, Patrice O.  
TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,049B  
FILING DATE: 09-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Charles K. Sholtz  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3  
INDIVIDUAL ISOLATE: region

US-08-240-049B-13  
Query Match 98.3%; Score 1076; DB 1; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.4e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPT 60  
DB 61 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPT 120  
QY 61 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYSGSTGPGVVSSTVTLVNVATGAQAVARS 120  
DB 121 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYSGSTGPGVVSSTVTLVNVATGAQAVARS 180  
QY 121 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
DB 181 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 240  
QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
DB 241 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 270

## RESULT 3

US-08-259-148A-15  
Sequence 15, Application US/08259148A  
Patent No. 5741490

## GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R.  
APPLICANT: Bradley, Daniel W.  
APPLICANT: Twu, Jr.-Shin  
APPLICANT: Purdy, Michael A.  
APPLICANT: Tam, Albert W.  
APPLICANT: Krawczynski, Krzysztof Z.  
APPLICANT: Yarbough, Patrice D.

TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/259,148A  
FILING DATE: 13-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 822,335  
FILING DATE: 17-JAN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 505,888  
FILING DATE: 05-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 420,921  
FILING DATE: 13-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 367,486  
FILING DATE: 16-JUN-1989

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 336,672  
FILING DATE: 11-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 208,997  
FILING DATE: 17-JUN-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0093.20  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 9

US-08-259-148A-15  
Query Match 98.3%; Score 1076; DB 1; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.4e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPT 60  
DB 61 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPT 120  
QY 61 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYSGSTGPGVVSSTVTLVNVATGAQAVARS 120  
DB 121 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYSGSTGPGVVSSTVTLVNVATGAQAVARS 180  
QY 121 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
DB 181 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 240  
QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
DB 241 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 270

Query Match 98.3%; Score 1076; DB 1; Length 327;

Best Local Similarity 100.0%; Pred. No. 4.4e-104;

Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPT 60  
DB 61 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPT 120  
QY 61 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYSGSTGPGVVSSTVTLVNVATGAQAVARS 120  
DB 121 PSPAPSRPFSVLRANDVWLSTAAEYDQSTYSGSTGPGVVSSTVTLVNVATGAQAVARS 180  
QY 121 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
DB 181 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 240  
QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
DB 241 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 270

7  
102

RESULT 4  
US-08-484-054-15  
; Sequence 15, Application US/08484054  
; Patent No. 520689  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.  
; APPLICANT: Bradley, Daniel W.  
; APPLICANT: Twu, Jr-Shin  
; APPLICANT: Purdy, Michael A.  
; APPLICANT: Tam, Albert W.  
; APPLICANT: Krawczynski, Krzysztof Z.  
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,054  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 870,985  
; FILING DATE: 20-APRIL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 822,335  
; FILING DATE: 17-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 505,888  
; FILING DATE: 05-APRIL-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 420,921  
; FILING DATE: 13-OCTOBER-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 367,486  
; FILING DATE: 16-JUNE-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 336,672  
; FILING DATE: 11-APRIL-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 208,997  
; FILING DATE: 17-JUNE-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 4600-0093.38  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 327 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 9  
US-08-484-054-15

Query Match 98.3%; Score 1076; DB 1; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.4e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLFYSRPVVSANGEPVTKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPT 60  
Db 61 QLFYSRPVVSANGEPVTKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPT 120  
Qy 61 PSPAPSPFVSRLRANDVWLSTLAAEYDQSTYGSSTGVPVYSDSVTLNNVATGAQAVARS 120  
Db 121 PSPAPSPFVSRLRANDVWLSTLAAEYDQSTYGSSTGVPVYSDSVTLNNVATGAQAVARS 180  
Qy 121 LDWTKVTLDRPLSTIQYSKTFFVLPLRGKLSFWEAGTTKAGYPYNTNTTASDQLLVEN 180  
Db 181 LDWTKVTLDRPLSTIQYSKTFFVLPLRGKLSFWEAGTTKAGYPYNTNTTASDQLLVEN 240  
Qy 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
Db 241 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 270

RESULT 5  
US-07-876-941A-15  
; Sequence 15, Application US/07876941A  
; Patent No. 5885768  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.  
; APPLICANT: Bradley, Daniel W.  
; APPLICANT: Tam, Albert W.  
; APPLICANT: Mitchell, Carl  
; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and  
; TITLE OF INVENTION: Antibodies  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/876,941A  
; FILING DATE: 01-MAY-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 822,335  
; FILING DATE: 17-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 505,888  
; FILING DATE: 05-APRIL-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 420,921  
; FILING DATE: 13-OCTOBER-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 367,486  
; FILING DATE: 16-JUNE-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 336,672  
; FILING DATE: 11-APRIL-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 208,997  
; FILING DATE: 17-JUNE-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 4600-0093.33  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 327 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 9  
US-07-876-941A-15

Query Match 98.3%; Score 1076; DB 2; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.4e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVWIQDYDNQHEQDRPT 60  
DB 61 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVWIQDYDNQHEQDRPT 120

QY 61 PSPAPSPFVILRANDVILWLSLTAAYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 120  
DB 121 PSPAPSPFVILRANDVILWLSLTAAYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 180

QY 121 LDWTKVTLDCRPLSTIQOYSKTFVPLRGLKSFWEAGTTKAGYPYNNYNTTASDQLLVEN 180  
DB 181 LDWTKVTLDCRPLSTIQOYSKTFVPLRGLKSFWEAGTTKAGYPYNNYNTTASDQLLVEN 240

QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
DB 241 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 270

RESULT 6  
US-08-542-634-17  
Sequence 17, Application US/08542634  
Patent No. 6214970  
GENERAL INFORMATION:  
APPLICANT: Fuerst, Thomas R.  
APPLICANT: McAtee, C. Patrick  
APPLICANT: Yarbough, Patrice O.  
APPLICANT: Zhang, Yifan  
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Ave., Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,634  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0293.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3

INDIVIDUAL ISOLATE: region  
US-08-542-634-17

Query Match 98.3%; Score 1076; DB 3; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.4e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVWIQDYDNQHEQDRPT 60  
DB 61 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGESRVWIQDYDNQHEQDRPT 120

QY 61 PSPAPSPFVILRANDVILWLSLTAAYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 120  
DB 121 PSPAPSPFVILRANDVILWLSLTAAYDQSTYGSSTGTPVVSVDVTLVNVATGAQAVARS 180

QY 121 LDWTKVTLDCRPLSTIQOYSKTFVPLRGLKSFWEAGTTKAGYPYNNYNTTASDQLLVEN 180  
DB 181 LDWTKVTLDCRPLSTIQOYSKTFVPLRGLKSFWEAGTTKAGYPYNNYNTTASDQLLVEN 240

QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
DB 241 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 270

RESULT 7  
US-08-477-292-1  
Sequence 17, Application US/08477292  
Patent No. 621641  
GENERAL INFORMATION:  
APPLICANT: Fuerst, Thomas  
APPLICANT: McAtee, Patrick  
APPLICANT: Yarbough, Patrice  
APPLICANT: Zhang, Yifan  
TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathleen M. Desjardins, M.D.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: US  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,292  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/327,952  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, Allan A  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: G32P5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)369-9500  
TELEFAX: (415)368-0709  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3  
INDIVIDUAL ISOLATE: region  
US-08-477-292-17

Query Match 98.3%; Score 1076; DB 3; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.4e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGPTVKLYTSVENAQQDKGIAIPHDIDLGSRVVIQDYNQHEQDRPT 60  
DB 61 QLFYSRPVVSANGPTVKLYTSVENAQQDKGIAIPHDIDLGSRVVIQDYNQHEQDRPT 120

QY 61 PSPAPSPFVSLRANDVWLSTAAEYDQSTGSGTGPVYVSDSVTLNVNATGAQAVARS 120  
DB 121 PSPAPSPFVSLRANDVWLSTAAEYDQSTGSGTGPVYVSDSVTLNVNATGAQAVARS 180

QY 121 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNNTTASDQLLVEN 180  
DB 181 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNNTTASDQLLVEN 240

QY 181 AAGHRVAISTVTTSLGAGPVSISAVAVLAP 210  
DB 241 AAGHRVAISTVTTSLGAGPVSISAVAVLAP 270

RESULT 8 / 985-15  
US-07-870-985A-15  
; Sequence 15 Application US/07870985A  
; Patent No. 4,554,92  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.  
; APPLICANT: Bradley, Daniel W.  
; APPLICANT: Two, Jr-Shin  
; APPLICANT: Purdy, Michael A.  
; APPLICANT: Tam, Albert W.  
; APPLICANT: Krawczynski, Krzysztof Z.  
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07870,985A  
; FILING DATE: 20-APRIL-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US 822,335  
; FILING DATE: 17-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 505,888  
; FILING DATE: 05-APRIL-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 420,921  
; FILING DATE: 13-OCTOBER-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 367,486  
; FILING DATE: 16-JUNE-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 336,672  
; FILING DATE: 11-APRIL-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 208,997  
; FILING DATE: 17-JUNE-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 4600-0093.30  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 9  
US-07-870-985A-15

Query Match 98.3%; Score 1076; DB 4; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.4e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGPTVKLYTSVENAQQDKGIAIPHDIDLGSRVVIQDYNQHEQDRPT 60  
DB 61 QLFYSRPVVSANGPTVKLYTSVENAQQDKGIAIPHDIDLGSRVVIQDYNQHEQDRPT 120

QY 61 PSPAPSPFVSLRANDVWLSTAAEYDQSTGSGTGPVYVSDSVTLNVNATGAQAVARS 120  
DB 121 PSPAPSPFVSLRANDVWLSTAAEYDQSTGSGTGPVYVSDSVTLNVNATGAQAVARS 180

QY 121 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNNTTASDQLLVEN 180  
DB 181 LDWTKVTLDRPLSTIQOYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNNTTASDQLLVEN 240

QY 181 AAGHRVAISTVTTSLGAGPVSISAVAVLAP 210  
DB 241 AAGHRVAISTVTTSLGAGPVSISAVAVLAP 270

RESULT 9 / 13703-17  
PCT-US95-13703-17  
; Sequence 17 Application PC/TUS9513703  
; GENERAL INFORMATION:  
; APPLICANT: HEPATITIS E VIRUS ANTIGENS AND  
; TITLE OF INVENTION: USES THEREFOR  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: P.O. Box 60850  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-0850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13703  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 4600-0293.41  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3  
INDIVIDUAL ISOLATE: region  
PCT-US95-13703-17

Query Match 98.3%; Score 1076; DB 5; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.4e-104; Indels 0; Gaps 0;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QLFYSRPVWSANGPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNOHQDRPT 60  
DB 61 QLFYSRPVWSANGPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNOHQDRPT 120  
QY 61 PPSAPSPFVSLRVANDVWLSTAAEYDQSTYGSSTGPTVVSVDVTLNVNATGAQAVARS 120  
DB 121 PPSAPSPFVSLRVANDVWLSTAAEYDQSTYGSSTGPTVVSVDVTLNVNATGAQAVARS 180  
QY 121 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
DB 181 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKSFWEAGTTKAGYPYNYNTTASDQLLVEN 240  
QY 181 AAGRVAISTVTTSLGAGPVSISAVAVLAP 210  
DB 241 AAGRVAISTVTTSLGAGPVSISAVAVLAP 270

RESULT 10  
US-08-259-148A-17  
Sequence 17 Application US/08259148A  
Patent No. 5741490  
GENERAL INFORMATION:  
APPLICANT: Reyes, Gregory R.  
APPLICANT: Bradley, Daniel W.  
APPLICANT: Twu, Jr-Shin  
APPLICANT: Purdy, Michael A.  
APPLICANT: Tam, Albert W.  
APPLICANT: Krawczynski, Krzysztof Z.  
APPLICANT: Yarbough, Patrice D.  
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/259,148A  
FILING DATE: 13-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: US 822,335  
FILING DATE: 17-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 505,888  
FILING DATE: 05-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 420,921  
FILING DATE: 13-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 367,486  
FILING DATE: 16-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 336,672  
FILING DATE: 11-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 208,997  
FILING DATE: 17-JUN-1988

ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0093.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 436 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9  
US-08-259-148A-17  
Query Match 98.3%; Score 1076; DB 1; Length 436;  
Best Local Similarity 100.0%; Pred. No. 6.7e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QLFYSRPVWSANGPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNOHQDRPT 60  
DB 170 QLFYSRPVWSANGPTVKLYTSVENAQQDKGIAIPHDIDLGESRVVIQDYDNOHQDRPT 229  
QY 61 PPSAPSPFVSLRVANDVWLSTAAEYDQSTYGSSTGPTVVSVDVTLNVNATGAQAVARS 120  
DB 230 PPSAPSPFVSLRVANDVWLSTAAEYDQSTYGSSTGPTVVSVDVTLNVNATGAQAVARS 289  
QY 121 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
DB 290 LDWTKVTLDRPLSTIQOYSKTFVPLRGLKSFWEAGTTKAGYPYNYNTTASDQLLVEN 349  
QY 181 AAGRVAISTVTTSLGAGPVSISAVAVLAP 210  
DB 350 AAGRVAISTVTTSLGAGPVSISAVAVLAP 379  
RESULT 11  
US-08-484-054-17  
Sequence 17 Application US/08484054  
Patent No. 5770689  
GENERAL INFORMATION:  
APPLICANT: Reyes, Gregory R.  
APPLICANT: Bradley, Daniel W.  
APPLICANT: Twu, Jr-Shin  
APPLICANT: Purdy, Michael A.  
APPLICANT: Tam, Albert W.  
APPLICANT: Krawczynski, Krzysztof Z.  
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,054  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: US 870,985  
FILING DATE: 20-APRIL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 822,335

FILING DATE: 17-JAN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 505,888  
 FILING DATE: 05-APRIL-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 420,921  
 FILING DATE: 13-OCTOBER-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 367,486  
 FILING DATE: 16-JUNE-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 336,672  
 FILING DATE: 11-APRIL-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 208,997  
 FILING DATE: 17-JUNE-1988  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sholtz, Charles K.  
 REGISTRATION NUMBER: 38,615  
 REFERENCE/DOCKET NUMBER: 4600-0093.38  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0880  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 436 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9  
 US-08-484-054-17

Query Match 98.3%; Score 1076; DB 1; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-104;  
 Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIALPHIDILGESRVWIDYDNOHQEDRPT 60  
 DB 170 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIALPHIDILGESRVWIDYDNOHQEDRPT 229  
 QY 61 PSPAPSRPFVLRANDVWLSTAAEYDQSTYGSSTGPTVYVSDSVTLNVNATGAQAVARS 120  
 DB 230 PSPAPSRPFVLRANDVWLSTAAEYDQSTYGSSTGPTVYVSDSVTLNVNATGAQAVARS 289  
 QY 121 LDWTKVTLDCRPLSTIQOYKTFPVLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
 DB 290 LDWTKVTLDCRPLSTIQOYKTFPVLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 349  
 QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
 DB 350 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 379

RESULT 12  
 US-07-876-941A-17  
 Sequence 17 Application US/07876941A  
 Patent No. 5985266  
 GENERAL INFORMATION:  
 APPLICANT: Reyes, Gregory R.  
 APPLICANT: Bradley, Daniel W.  
 APPLICANT: Tam, Albert W.  
 APPLICANT: Mitchell, Carl  
 TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and  
 TITLE OF INVENTION: Antibodies  
 NUMBER OF SEQUENCES: 76  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Dehlinger & Associates  
 STREET: 350 Cambridge Avenue, Suite 250  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA

ZIP: 94306  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/876,941A  
 FILING DATE: 01-MAY-1992  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 822,335  
 FILING DATE: 17-JAN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 505,888  
 FILING DATE: 05-APRIL-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 420,921  
 FILING DATE: 13-OCTOBER-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 367,486  
 FILING DATE: 16-JUNE-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 336,672  
 FILING DATE: 11-APRIL-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 208,997  
 FILING DATE: 17-JUNE-1988  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sholtz, Charles K.  
 REGISTRATION NUMBER: 38,615  
 REFERENCE/DOCKET NUMBER: 4600-0093.33  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0880  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 436 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9  
 US-07-876-941A-17

Query Match 98.3%; Score 1076; DB 2; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-104;  
 Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIALPHIDILGESRVWIDYDNOHQEDRPT 60  
 DB 170 QLFYSRPVVSANGEPTVKLYTSVENAQQDKGIALPHIDILGESRVWIDYDNOHQEDRPT 229  
 QY 61 PSPAPSRPFVLRANDVWLSTAAEYDQSTYGSSTGPTVYVSDSVTLNVNATGAQAVARS 120  
 DB 230 PSPAPSRPFVLRANDVWLSTAAEYDQSTYGSSTGPTVYVSDSVTLNVNATGAQAVARS 289  
 QY 121 LDWTKVTLDCRPLSTIQOYKTFPVLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
 DB 290 LDWTKVTLDCRPLSTIQOYKTFPVLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 349  
 QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
 DB 350 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 379

RESULT 13  
 US-07-870-985A-17  
 Sequence 17 Application US/07870985A  
 Patent No. 6455492  
 GENERAL INFORMATION:  
 APPLICANT: Reyes, Gregory R.

; APPLICANT: Bradley, Daniel W.  
 ; APPLICANT: Twu, Jr-Shin  
 ; APPLICANT: Purdy, Michael A.  
 ; APPLICANT: Tam, Albert W.  
 ; APPLICANT: Krawczynski, Krzysztof Z.  
 ; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
 ; NUMBER OF SEQUENCES: 22  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dehlinger & Associates  
 ; STREET: 350 Cambridge Avenue, Suite 250  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94306  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/870,985A  
 ; FILING DATE: 20-APRIL-1992  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 822,335  
 ; FILING DATE: 17-JAN-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 505,888  
 ; FILING DATE: 05-APRIL-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 420,921  
 ; FILING DATE: 13-OCTOBER-1989  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 367,486  
 ; FILING DATE: 16-JUNE-1989  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 336,672  
 ; FILING DATE: 11-APRIL-1989  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 208,997  
 ; FILING DATE: 17-JUNE-1988  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fabian, Gary R.  
 ; REGISTRATION NUMBER: 33,875  
 ; REFERENCE/DOCKET NUMBER: 4600-0093.30  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 324-0960  
 ; TELEFAX: (415) 324-0960  
 ; INFORMATION FOR SEQ ID NO: 17:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 436 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ORIGINAL SOURCE:  
 ; INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 9  
 ; US-07-870-985A-17  
 ;  
 ; Query Match 98.3%; Score 1076; DB 4; Length 436;  
 ; Best Local Similarity 100.0%; Pred. No. 6.7e-104;  
 ; Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 ;  
 QY 1 QLFYSPVVSANGEPTVKLYTSVENAQDQKGAIPHDIDLGSRVVIQDYDNQHEQDRPT 60  
 DB 170 QLFYSPVVSANGEPTVKLYTSVENAQDQKGAIPHDIDLGSRVVIQDYDNQHEQDRPT 229  
 QY 61 PSPAPSRPFSVLRANDVLSLTAAYDQSTYGSSTGPGVYSDSVTLNVNATGAQAVARS 120  
 DB 230 PSPAPSRPFSVLRANDVLSLTAAYDQSTYGSSTGPGVYSDSVTLNVNATGAQAVARS 289  
 QY 121 LDWTKVTLDRPLSTIQYSKTFFVFLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180

DB 290 LDWTKVTLDRPLSTIQYSKTFFVFLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 349  
 QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
 DB 350 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 379  
 ;  
 RESULT 14  
 ; US-08-542-634-27  
 ; Sequence 27, Application US/08542634  
 ; Patent No. 6214970  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fuerst, Thomas R.  
 ; APPLICANT: McAtee, C. Patrick  
 ; APPLICANT: Yarbough, Patrice O.  
 ; APPLICANT: Zhang, Yifan  
 ; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dehlinger & Associates  
 ; STREET: 350 Cambridge Ave., Suite 250  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94306  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/542,634  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fabian, Gary R.  
 ; REGISTRATION NUMBER: 33,875  
 ; REFERENCE/DOCKET NUMBER: 4600-0293.30  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 324-0880  
 ; TELEFAX: (415) 324-0960  
 ; INFORMATION FOR SEQ ID NO: 27:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 525 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ORIGINAL SOURCE:  
 ; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)  
 ; INDIVIDUAL ISOLATE: r62kDa, 56.5 kDa  
 ; US-08-542-634-27  
 ;  
 ; Query Match 98.3%; Score 1076; DB 3; Length 525;  
 ; Best Local Similarity 100.0%; Pred. No. 8.8e-104;  
 ; Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 ;  
 QY 1 QLFYSPVVSANGEPTVKLYTSVENAQDQKGAIPHDIDLGSRVVIQDYDNQHEQDRPT 60  
 DB 283 QLFYSPVVSANGEPTVKLYTSVENAQDQKGAIPHDIDLGSRVVIQDYDNQHEQDRPT 342  
 QY 61 PSPAPSRPFSVLRANDVLSLTAAYDQSTYGSSTGPGVYSDSVTLNVNATGAQAVARS 120  
 DB 343 PSPAPSRPFSVLRANDVLSLTAAYDQSTYGSSTGPGVYSDSVTLNVNATGAQAVARS 402  
 QY 121 LDWTKVTLDRPLSTIQYSKTFFVFLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 180  
 DB 403 LDWTKVTLDRPLSTIQYSKTFFVFLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN 462  
 QY 181 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 210  
 DB 463 AAGHRVAISTYTTSLGAGPVSISAVAVLAP 492



RESULT 15  
PCT-US95-13703-27  
; Sequence 27, Application PC/TUS9513703  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND  
; TITLE OF INVENTION: USES THEREFOR  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: P.O. Box 60850  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-0850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13703  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 4600-0293.41  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 525 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma strain)  
; INDIVIDUAL ISOLATE: r62kDa, 56.5 kDa  
PCT-US95-13703-27

Query Match 98.3%; Score 1076; DB 5; Length 525;  
Best Local Similarity 100.0%; Pred. No. 8.8e-104;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	QLFYSRPVVSANGPTVKLYTSVENAQODKGIAIPHDIDLGESRVVIQDYDNOHQDRPT	60
Db	283	QLFYSRPVVSANGPTVKLYTSVENAQODKGIAIPHDIDLGESRVVIQDYDNOHQDRPT	342
Qy	61	PSPAPSRPFSVLRANDVLMVLSLTAAYDQSTYGSSTGTVVSDSVTLVNVATGAQAVRS	120
Db	343	PSPAPSRPFSVLRANDVLMVLSLTAAYDQSTYGSSTGTVVSDSVTLVNVATGAQAVRS	402
Qy	121	LDWTKVTLDGRLPTIQOYSKTFVPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN	180
Db	403	LDWTKVTLDGRLPTIQOYSKTFVPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVEN	462
Qy	181	AAGHRVAISTYTTSLGAGPVSISAVAVLAP	210
Db	463	AAGHRVAISTYTTSLGAGPVSISAVAVLAP	492

Search completed: April 28, 2004, 14:22:28  
Job time : 23 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2004, 14:16:38 ; Search time 18 Seconds  
(without alignments)  
616.164 Million cell updates/sec

Title: US-10-089-292A-2

Perfect score: 1095

Sequence: 1 QLFYSRPVVSANGSPVTKLY.....SLGAGPVISAVAVLAPPPR 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1076	98.3	660	1 VST2 HEVBU	P29326 hepatitis e
2	1076	98.3	660	1 VST2 HEVPA	P33426 hepatitis e
3	1060	96.8	485	1 VST2 HEVRH	Q00270 hepatitis e
4	1060	96.8	660	1 VST2 HEVWY	Q04611 hepatitis e
5	1043	95.3	659	1 VST2 HEVME	Q03500 hepatitis e
6	91.5	8.4	1305	1 FTSK YERPE	Q829C7 yersinia pe
7	90	8.2	1045	1 GUNB CELFI	P26225 cellulomona
8	89	8.1	477	1 KPYK BORBU	OS1323 borrelia bu
9	89	8.1	1802	1 HKR1 YEAST	P41809 saccharomyc
10	84.5	7.7	431	1 PURA ECOLI	P12283 escherichia
11	84.5	7.7	431	1 PURA SALTY	Q8XGPI salmonella
12	84	7.7	432	1 ENO BIFLO	Q8G519 bifidobacte
13	83	7.6	617	1 FXK1 MOUSE	P42128 mus musculu
14	83	7.6	765	1 EGLX ECOLI	P33363 escherichia
15	82.5	7.5	234	1 POLN RHDV3	P27411 rabbit hemo
16	82.5	7.5	278	1 EUTJ ECOLI	P77277 escherichia
17	82	7.5	335	1 GCAB MOUSE	P01864 mus musculu
18	81.5	7.4	2344	1 POLN RHDV	P27410 rabbit hemo
19	81.5	7.4	3354	1 CADN HUMAN	Q9h251 homo sapien
20	81	7.4	418	1 EFTU CHLRE	P17256 chlamydomon
21	81	7.4	563	1 GUNB CLOTRM	P04956 clostridium
22	80.5	7.4	799	1 AFESK STROCO	P54741 streptomyce
23	79.5	7.3	279	1 EUTJ SALTY	P41794 salmonella
24	79.5	7.3	1020	1 CARY LACPL	Q9rls9 lactobacill
25	79	7.2	541	1 YD18 MYCTU	Q10631 mycobacteri
26	78.5	7.2	335	1 ISH1 MYCTU	OS3458 mycobacteri
27	78.5	7.2	584	1 FRAC BORPE	P35077 bordetella
28	78.5	7.2	1256	1 ATL STAAU	P52081 staphylococ
29	78	7.1	569	1 YICH ECOLI	P31433 escherichia
30	78	7.1	983	1 Y142 HUMAN	Q14157 homo sapien
31	78	7.1	1077	1 E2K3 CAEL	Q19192 caenorhabdi
32	77	7.0	308	1 XERD BIFLO	Q7zapl bifidobacte
33	77	7.0	433	1 KCG3 RAT	Q8r523 rattus norv

ALIGNMENTS

RESULT 1									
VST2	HEVBU	STANDARD;	PRT;	660 AA.					
ID	VST2 HEVBU								
AC	P29326;								
DT	01-DEC-1992 (Rel. 24, Created)								
DT	01-DEC-1992 (Rel. 24, Last sequence update)								
DT	01-FEB-1994 (Rel. 28, Last annotation update)								
DE	Structural protein 2 precursor (ORF2).								
OS	Hepatitis E virus (strain Burma) (HEV).								
OC	Viruses; ssRNA positive-strand viruses, no DNA stage;								
OC	Hepatitis E-like viruses.								
OX	NCBI_TaxID=31767;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=92024067; PubMed=1926770;								
RA	Tam A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,								
RA	Fry K.E., Reyes G.R.;								
RT	"Hepatitis E virus (HEV): molecular cloning and sequencing of the								
RT	full-length viral genome.;"								
RL	Virology 185:120-131(1991).								
CC	-1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING								
CC	THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA								
CC	BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>								
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).								
CC	-----								
DR	EMBL; M73218; AAA45736.1; -								
DR	PIR; C40778; VHWWH2.								
DR	InterPro; IPR004261; SP2.								
DR	InterPro; IPR008975; Viral_cap_coat.								
DR	Pfam; PF03014; SP2; 1.								
KW	Signal.								
FT	SIGNAL	1	19	POTENTIAL.					
FT	CHAIN	20	660	STRUCTURAL PROTEIN 2.					
SQ	SEQUENCE	660 AA;	70978 MW;	5832A013CCC4A61C CRC64;					
Query Match									
Best Local Similarity 100.0%; Pred. No. 6.2e-86;									
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
-----									
Qy	1	QLFYSRPVVSANGSPVTKLYTSVENAQDKGIALPHDIDLGESRVVIQDYDNQHQEDRPT	60						
Db	394	QLFYSRPVVSANGSPVTKLYTSVENAQDKGIALPHDIDLGESRVVIQDYDNQHQEDRPT	453						
Qy	61	PSPAPSPFVSLRANDVLSLTAAYDQSTGYSGSTGPFVYVSDSVTLNVNATGAQAVRS	120						
Db	454	PSPAPSPFVSLRANDVLSLTAAYDQSTGYSGSTGPFVYVSDSVTLNVNATGAQAVRS	513						
Qy	121	LDWTKVTLDGRPLSTIQOYKSTFFVLPRLKGLSFWBAGTTKAGYPYNYNTTASDQLLVEN	180						

P24153 vibrio chol  
P98168 homo sapien  
P03109 bovine papi  
P74104 synechocyst  
P38479 saccharomyc  
Q64760 avian adeno  
P46821 homo sapien  
P50378 gonium pect  
Q92qt4 rhizobium m  
Q09825 schizosacch  
P30601 exophiala d

Db 514 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVEN 573  
 QY 181 AAGRVAISTYTTSLGAGPVSISAVAVLAP 210  
 Db 574 AAGRVAISTYTTSLGAGPVSISAVAVLAP 603

## RESULT 2

VST2 HEVPA STANDARD; PRT; 660 AA.  
 ID VST2 HEVPA AC P33426;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE Structural protein 2 precursor (ORF2).  
 OS Hepatitis E virus (strain Pakistan) (HEV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;  
 OC Hepatitis E-like viruses.  
 'OX NCBI\_TaxID=33774;  
 RN NCBI\_TaxID=33774;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92115700; PubMed=1731327;  
 RA Tearev S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J.,  
 RA Malik I.A., Iqbal M., Purcell R.H.;  
 RT "Characterization of a prototype strain of hepatitis E virus."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).  
 CC -1- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING  
 CC THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA  
 CC BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; M80581; AAA45727.1; --  
 DR InterPro; IPR004261; SP2.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF03014; SP2; 1.  
 FT CHAIN 1 22 BY SIMILARITY.  
 FT SIGNAL 23 660 STRUCTURAL PROTEIN 2.  
 FT CHAIN 23 660  
 FT SEQUENCE 660 AA; 70980 MW; 8085BC53CFB46FD3 CRC64;

Query Match 98.3%; Score 1076; DB 1; Length 660;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-86;  
 Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGEPVTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHQDRPT 60  
 Db 394 QLFYSRPVVSANGEPVTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHQDRPT 453  
 QY 61 PSPAPSRPFSVLRANDVWLSTLTAAYDQSTYSGSTGTPVVSVDVTLVNVATGAQAVARS 120  
 Db 454 PSPAPSRPFSVLRANDVWLSTLTAAYDQSTYSGSTGTPVVSVDVTLVNVATGAQAVARS 513  
 QY 121 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVEN 180  
 Db 514 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVEN 573  
 QY 181 AAGRVAISTYTTSLGAGPVSISAVAVLAP 210  
 Db 574 AAGRVAISTYTTSLGAGPVSISAVAVLAP 603

## RESULT 3

VST2 HEVRH STANDARD; PRT; 485 AA.  
 ID VST2 HEVRH AC Q00270;  
 DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DE Structural protein 2 (Fragment).  
 OS Hepatitis E virus (isolate Rhesus) (HEV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;  
 OC Hepatitis E-like viruses.  
 'OX NCBI\_TaxID=31766;  
 RN NCBI\_TaxID=31766;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92261377; PubMed=1594074;  
 RA Uchida T., Suzuki K., Hayashi N., Iida F., Hara T., Oo S.S.,  
 RA Wang C.-K., Shikata T., Ichikawa M., Rikihisa T., Mizuno K.,  
 RA Win K.M.;  
 RT "Hepatitis E virus: cDNA cloning and expression."  
 RL Microbiol. Immunol. 36:67-79(1992).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; D90274; BAA20910.1; --  
 DR InterPro; IPR004261; SP2.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF03014; SP2; 1.  
 FT NON\_TER 1 485  
 FT NON\_TER 485  
 FT SEQUENCE 485 AA; 52317 MW; 5A4FD9273AC74F94 CRC64;

Query Match 96.8%; Score 1060; DB 1; Length 485;  
 Best Local Similarity 98.6%; Pred. No. 1e-84;  
 Matches 207; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QLFYSRPVVSANGEPVTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHQDRPT 60  
 Db 262 QLFYSRPVVSANGEPVTVKLYTSVENAQDQKGIAPHDIDLGESRVVIQDYDNOHQDRPT 321  
 QY 61 PSPAPSRPFSVLRANDVWLSTLTAAYDQSTYSGSTGTPVVSVDVTLVNVATGAQAVARS 120  
 Db 322 PSPAPSRPFSVLRANDVWLSTLTAAYDQSTYSGSTGTPVVSVDVTLVNVATGAQAVARS 381  
 QY 121 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVEN 180  
 Db 382 LDWTKVTLDCRPLSTIQYKSTFFVPLRGLKLSFWEAGTTKAGYPYNNYNTTASDQLLVEN 441  
 QY 181 AAGRVAISTYTTSLGAGPVSISAVAVLAP 210  
 Db 442 AAGRVAISTYTTSLGAGPVSISAVAVLAP 471

## RESULT 4

VST2 HEVMY STANDARD; PRT; 660 AA.  
 ID VST2 HEVMY AC Q04611;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Structural protein 2 precursor (ORF2).  
 OS Hepatitis E virus (strain Myanmar) (HEV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage;  
 OC Hepatitis E-like viruses.  
 'OX NCBI\_TaxID=31769;  
 RN NCBI\_TaxID=31769;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93227573; PubMed=8470371;  
 RA Aye T.T., Uchida T., Ma M.Z., Iida F., Shikata T., Ichikawa M.,  
 RA Rikihisa T., Winn K.;  
 RT "Sequence and gene structure of the hepatitis E virus isolated from  
 RT Myanmar."  
 RL Virus Genes 7:95-109(1993).

[illegible]

CC dependence of magnetic flux is a function of it and negative generating positive supercoils in front of it and negative

**All rights reserved.**

CC dependence of magnetic flux is a function of it and negative generating positive supercoils in front of it and negative



Query Match 8.2%; Score 90; DB 1; Length 1045;  
 Best Local Similarity 26.3%; Pred. No. 6.7;  
 Matches 65; Conservative 18; Mismatches 100; Indels 64; Gaps 11;

QY 11 ANGSEPTVKLYTSVENAQODKGIAPHIDLDGESRVLIQDY-----DNOHQD---RPTSP 63  
 DB 544 AASDVTLISANYSECAQSGKVSAGTILGYVELSCVQDHPGQSQRRRIQRLT-GP 602  
 QY 64 APSRPFVSLRANVWLWLSLTAEE-----YDQSTYSGSTGPVYVSDSVTL-----VN 109  
 DB 603 AGWNP-----ANDPSYTLGLTQALAKASAITLYDGLSTLVWGKEFTGTTTTPPTPTGP 657  
 QY 110 VATGAQAVARSLDWTKVT-----LDGRPLSTIQYKSTFFVLPLRGKLSFWEAGTTKA-- 162  
 DB 658 VATGVTTVTTCASLSWAASDAGSGVAGYELXVQGTQTTL-----VGTITAAA 704  
 QY 163 -----GYPNYNTTASDQLLVENAGHRVAJSTVTSLGAGPVS-----ISAV----- 205  
 DB 705 YILRDLTPGTATYVVKAKVAGNVSAASAAVTFTTDTTGTEPTTPTGTPVASAVTSTG 764  
 QY 206 AVLAPPP 212  
 DB 765 ATLAWAP 771

RESULT 8  
 KPYK BORBU  
 ID KPYK BORBU STANDARD; PRT; 477 AA.  
 AC 051323;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Pyruvate kinase (EC 2.7.1.40) (PK).  
 GN PYK OR BB0348.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 OC NCBI\_TaxID=139;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE=98065943; PubMed=9403685;  
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
 RA Peterson J., Kervilave A.R., Quackenbush J., Salzberg S., Hanson M.,  
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
 RA Uterback T., Watthey L., McDonald L., Artiach P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 CC -!- CATALYTIC ACTIVITY: ATP + pyruvate = ADP + phosphoenolpyruvate.  
 CC -!- COFACTOR: Requires magnesium and potassium.  
 CC -!- PATHWAY: Glycolysis; final step.  
 CC -!- SUBUNIT: Homotrimer (By similarity).  
 CC -!- SIMILARITY: Belongs to the pyruvate kinase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; AE001141; AAC66733.1; -.  
 DR PIR; C70143; C70143.  
 DR HSP; P14178; 150T.  
 DR TIGR; BB0348; -.  
 DR InterPro; IPR001697; Pyruvate\_kinase.  
 DR Pfam; PF00224; PK; 1.  
 DR Pfam; PF02887; PK\_C; 1.

DR PRINTS; PR01050; PYRUVTKINASE.  
 DR ProDom; PD001009; Pyruvate kinase; 1.  
 DR TIGRFAMs; TIGR01064; pyruv kin; 1.  
 DR PROSITE; PS00110; PYRUVATE KINASE; 1.  
 KW Pyruvate; Transferase; Kinase; Glycolysis; Magnesium;  
 KW Complete proteome. 214 BY SIMILARITY.  
 FT ACT\_SITE 214 216 MAGNESIUM (BY SIMILARITY).  
 FT METAL 216 216 MAGNESIUM (BY SIMILARITY).  
 FT METAL 237 237 MAGNESIUM (BY SIMILARITY).  
 FT METAL 238 238 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 477 AA; 53032 MW; 40D3CCAC9ACB7D98 CRC64;

Query Match 8.1%; Score 89; DB 1; Length 477;  
 Best Local Similarity 22.2%; Pred. No. 3;  
 Matches 51; Conservative 35; Mismatches 74; Indels 70; Gaps 11;

QY 8 VVSANGSEPTVKLYTSVEN-----AQDKGIAIP-HDIDGESRV-- 45  
 DB 200 ILTASGNPDVKLIISKIENQEGIDNIEIAKASYGIMVARGDMGVEIPEADVPFIAQLKITQ 259  
 QY 46 -----VIQDYDNQH---EQDRPTSPAPSRPFSVLRANDVWLWLSLTAAYEQSTYGS 94  
 DB 260 TCIKYGIPVITATQMLHTMIENPRPTAEVSDIANAILNGTDAILMS-----GETAYGK 313  
 QY 95 STGPEVYVSDSVTLVNVATGAQAVARSLDWT-----KVTLDRPLSTI 136  
 DB 314 -----YPIEAVKM--MTSIAKVEKHKWTLTKDELFFDKSITRNYIKCAIDATKLMDI 366  
 QY 137 QQYKSTFFVLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAGHRV 186  
 DB 367 -----KAIIVDSLKGKTARIMA--TYRASVPL-FITNSERLARELALSYGV 410

RESULT 9  
 HKR1\_YEAST  
 ID HKR1\_YEAST STANDARD; PRT; 1802 AA.  
 AC P41809;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hansenula MRK11 killer toxin-resistant protein 1 precursor.  
 GN HKR1 OR YDR420W.  
 OC Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OC NCBI\_TaxID=4932;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=YNN 295;  
 RX MEDLINE=94156857; PubMed=8113191;  
 RA Kasahara S., Yamada H., Mio T., Shiratori Y., Miyamoto C.,  
 RA Yabe T., Nakajima T., Ichishima E., Furuichi Y.,  
 RT "Cloning of the Saccharomyces cerevisiae gene whose overexpression  
 RT overcomes the effects of HM-1 killer toxin, which inhibits  
 RT beta-glucan synthesis.";  
 RT J. Bacteriol. 176:1488-1499(1994).  
 CC -!- FUNCTION: Could regulate beta-glucan synthesis. Overexpression  
 CC provides resistance to HM-1 killer toxin.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).  
 CC -!- PTM: Could be O-glycosylated in the serine/threonine-rich domain.  
 CC -!- SIMILARITY: SOME, TO YEAST MSB2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; S69101; AAB30051.1; -.  
 DR GenBank; 140912; -.  
 DR SGD; S0002828; HKR1.

KW Glycoprotein; Transmembrane; Repeat; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 1802  
FT FT HANSENULA MRAKII KILLER TOXIN-RESISTANT  
FT FT PROTEIN 1.  
FT TRANSSEM 1486 1506  
FT DOMAIN 23 1478  
FT DOMAIN 453 788  
FT FT  
FT FT  
FT REPEAT 453 480  
FT REPEAT 481 508  
FT REPEAT 509 536  
FT REPEAT 537 564  
FT REPEAT 565 592  
FT REPEAT 593 620  
FT REPEAT 621 648  
FT REPEAT 649 676  
FT REPEAT 677 704  
FT REPEAT 705 732  
FT REPEAT 733 760  
FT REPEAT 761 788  
FT CARBOHYD 24 24  
FT CARBOHYD 1252 1252  
FT CARBOHYD 1293 1293  
FT CARBOHYD 1342 1342  
FT CARBOHYD 1400 1400  
SQ SEQUENCE 1802 AA; E344CA6469785A24 CRC64;  
Query Match 8.1%; Score 89; DB 1; Length 1802;  
Best Local Similarity 26.3%; Pred. No. 17;  
Matches 44; Conservative 24; Mismatches 47; Indels 52; Gaps 8;  
QY 60 TPSPAPSPFVLRANDVLSLTAAEYDQSTGYSGTGPVVSDFVNVVATGAQAVAR 119  
DB 1200 TPYSP-----NSYMLP-TALIVSESTGPTTASFNPSTGSLPNAIEPAVASE 1249  
QY 120 SLDWKVT-----LDGRPLSTIQYKTFVFLPLRGLKSLFWEAGTTKAGYPNY 168  
DB 1250 PINHTLITGFTALNVFLVQNPSSAQ-----IFNELPLVK-----YPF 1291  
QY 169 NTTASDQLLVENAGHRVASTYTTSLGAG-----PVTSISAVAVL 208  
DB 1292 SNTSSE---LDNSTGE---LSTFILSYRSGSSTTLSPKSISSLSVV 1332  
RESULT 10  
PURA\_ECOLI  
ID PURA\_ECOLI STANDARD; PRT; 431 AA.  
AC P12283;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)  
DE (AdSS) (AMPSase).  
GN PURA OR ADEK OR B4177 OR C5261 OR Z5784 OR ECSS5153.  
OS Escherichia coli,  
OS Escherichia coli O6, and  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 217992, 83334;  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.  
RC STRAIN=K12;  
RX MEDLINE=89066719; PubMed=3058695;  
RA Wolfe S.A., Smith J.M.;  
RT "Nucleotide sequence and analysis of the pura gene encoding  
RT adenylosuccinate synthetase of Escherichia coli K12.";  
RL J. Biol. Chem. 263:19147-19153(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=95334362; PubMed=7610040;

RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,  
RA Blattner F.R.;  
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the  
RT region from 92.8 through 100 minutes.";  
RL Nucleic Acids Res. 23:2105-2119(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
RT of uropathogenic Escherichia coli.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potanousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
RN [6]  
RP PARTIAL SEQUENCE OF 1-9.  
RC STRAIN=K12 / EWG2;  
RX MEDLINE=97443975; PubMed=9298646;  
RA Link A.J., Robinson K., Church G.M.;  
RT "Comparing the predicted and observed properties of proteins encoded  
RT in the genome of Escherichia coli K-12.";  
RL Electrophoresis 18:1259-1313(1997).  
RN [7]  
RP ACTIVE SITE LYS-140.  
RX MEDLINE=90202896; PubMed=2108156;  
RA Dong Q., Fromm H.J.;  
RT "Chemical modification of adenylosuccinate synthetase from  
RT Escherichia coli by pyridoxal 5'-phosphate. Identification of an  
RT active site lysyl residue.";  
RL J. Biol. Chem. 265:6235-6240(1990).  
RN [8]  
RP SEQUENCE OF 145-147, AND MUTAGENESIS OF ARG-147.  
RX MEDLINE=91286237; PubMed=2061308;  
RA Dong Q., Liu F., Myers A.M., Fromm H.J.;  
RT "Evidence for an arginine residue at the substrate binding site of  
RT Escherichia coli adenylosuccinate synthetase as studied by chemical  
RT modification and site-directed mutagenesis.";  
RL J. Biol. Chem. 266:12228-12233(1991).  
RN [9]  
RP MUTAGENESIS.  
RX MEDLINE=92129317; PubMed=1733940;  
RA Liu F., Dong Q., Fromm H.J.;  
RT "Site-directed mutagenesis of the phosphate-binding consensus  
RT sequence in Escherichia coli adenylosuccinate synthetase.";  
RL J. Biol. Chem. 267:2398-2392(1992).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND REVISION TO 416.  
RX MEDLINE=96095802; PubMed=7490761;



RA Silva M.W., Poland B.W., Hoffman C.R., Fromm H.J., Honzatzko R.B.;  
 RT "Refined crystal structures of unligated adenylosuccinate synthetase  
 RL from *Escherichia coli*.";  
 RN J. Mol. Biol. 254:431-446(1995).  
 RP [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RX MEDLINE=97121272; PubMed=8961938;  
 RA Poland B.W., Lee S.F., Subramanian M.V., Siehl D.L., Anderson R.J.,  
 RA Fromm H.J., Honzatzko R.B.;  
 RT "Refined crystal structure of adenylosuccinate synthetase from  
 RT *Escherichia coli* complexed with hydantocidin 5'-phosphate, GDP,  
 RT HPO4(2-), Mg2+, and hadacidin.";  
 RL Biochemistry 35:15753-15759(1996).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=97153337; PubMed=9000627;  
 RA Poland B.W., Fromm H.J., Honzatzko R.B.;  
 RT "Crystal structures of adenylosuccinate synthetase from *Escherichia*  
 RT *coli* complexed with GDP, IMP hadacidin, NO3-, and Mg2+.";  
 RL J. Mol. Biol. 264:1013-1027(1996).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=99278141; PubMed=10346917;  
 RA Choe J.Y., Poland B.W., Fromm H.J., Honzatzko R.B.;  
 RT "Mechanistic implications from crystalline complexes of wild-type and  
 RT mutant adenylosuccinate synthetases from *Escherichia coli*.";  
 RL Biochemistry 38:6953-6961(1999).  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=9929706; PubMed=10364182;  
 RA Hou Z., Cashel M., Fromm H.J., Honzatzko R.B.;  
 RT "Effectors of the stringent response target the active site of  
 RT *Escherichia coli* adenylosuccinate synthetase.";  
 RL J. Biol. Chem. 274:17505-17510(1999).  
 CC -I- FUNCTION: Plays an important role in the de novo pathway of purine  
 CC nucleotide biosynthesis.  
 CC -I- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +  
 CC adenylosuccinate.  
 CC -I- COFACTOR: Binds 1 magnesium ion per subunit.  
 CC -I- PATHWAY: AMP biosynthesis; first committed step.  
 CC -I- SUBUNIT: Homodimer.  
 CC -I- SIMILARITY: Belongs to the adenylosuccinate synthetase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; J04199; AAA24446.1; -;  
 DR EMBL; U14003; AAA97073.1; -;  
 DR EMBL; AE000490; AAC77134.1; -;  
 DR EMBL; AE016771; AAN83683.1; -;  
 DR EMBL; AE005651; AAC59373.1; -;  
 DR EMBL; AP025568; BAB38576.1; -;  
 DR PIR; A98273; A98273.  
 DR PIR; S56402; AJECDS.  
 DR PDB; 1ADE; 26-JAN-96.  
 DR PDB; 1ADI; 10-JUN-96.  
 DR PDB; 1GIM; 23-DEC-96.  
 DR PDB; 1GIN; 12-FEB-97.  
 DR PDB; 1HON; 08-NOV-96.  
 DR PDB; 1HOP; 12-FEB-97.  
 DR PDB; 1KSH; 08-NOV-96.  
 DR PDB; 1KSH; 08-OCT-97.  
 DR PDB; 1NHT; 08-OCT-97.  
 DR PDB; 1SON; 04-SEP-97.  
 DR PDB; 1SOO; 04-SEP-97.  
 DR PDB; 1JUY; 24-JUN-97.  
 DR PDB; 1QF4; 02-DEC-99.

DR PDB; 1QF5; 02-DEC-99.  
 DR PDB; 1CG0; 17-JUN-99.  
 DR PDB; 1CG1; 17-JUN-99.  
 DR PDB; 1CG3; 17-JUN-99.  
 DR PDB; 1CG4; 17-JUN-99.  
 DR PDB; 1CH8; 29-DEC-99.  
 DR PDB; 1CIB; 05-APR-00.  
 DR PDB; 1KJX; 20-MAR-02.  
 DR PDB; 1KKE; 20-MAR-02.  
 DR PDB; 1KTF; 20-MAR-02.  
 DR SWISS-2DPAGE; P12283; COLI.  
 DR EcoGene; EG10790; purA.  
 DR HAMAP; MF\_00011; -; 1.  
 DR InterPro; IPR001114; Asucc\_synthase.  
 DR Pfam; PF00709; Adenylsucc synt; 1.  
 DR ProDom; PD001188; Asucc synthase; 1.  
 DR TIGRFAMs; TIGR00184; purA; 1.  
 DR PROSITE; PS00513; ADENYLOSUCCIN SYN 2; 1.  
 DR PROSITE; PS01266; ADENYLOSUCCIN SYN 1; 1.  
 KW Purine biosynthesis; Ligase; GTP-binding; Metal-binding; Magnesium;  
 KW 3D-structure; Complete proteome.  
 FT INIT MET 0  
 FT NP\_BIND 12 18 GTP (POTENTIAL).  
 FT ACT\_SITE 140 140  
 FT ACT\_SITE 147 147 PROBABLE.  
 FT METAL 13 13 MAGNESIUM.  
 FT METAL 40 40 MAGNESIUM (VIA CARBONYL OXYGEN).  
 FT MUTAGEN 12 12 G->V: SIGNIFICANT REDUCTION IN ACTIVITY.  
 Query Match 7.7%; Score 84.5; DB 1; Length 431;  
 Best Local Similarity 24.3%; Pred. No. 6.5;  
 Matches 51; Conservative 29; Mismatches 73; Indels 57; Gaps 11;  
 Qy 28 QDKGIALPHDIDGESRVVIQDY-----DNQHEQDRPTPS-----PAPSRPFVS---L 72  
 Db 87 EDGIPVRERLLSEACPLDILYHVALDNAREKARGAKAIGTTGRGIGPAYEDKVARGL 146  
 Qy 73 RANDVLWLSLTAAYEQDSTYSGSTGPVYVSDSVTLNVNATGAQAVARSLDWTVTLDGRP 132  
 Db 147 RVGDL-----FDKETAEKLEKWEYHFNQVNYK-----AEADVQKVLDDTMA 192  
 Qy 133 LSTI-----QQYSKTFVL--PLRGKLSFEAGTTKAGYYP--NYNTTASDQ 175  
 Db 193 VADILTSMVVDVSDLLDQARQGRGDFVFEQAQGLTLDIDHGT-----YPYVTSNTTAG- 247  
 Qy 176 LLVENAAGHVA-----ISTYTTSLGAGP 199  
 Db 248 VATSGGLGPRYVDYVLGILKAYSTRVGAGP 277  
 RESULT 11  
 PURA\_SALTY  
 ID PURA\_SALTY STANDARD; PRT; 431 AA.  
 AC QXKGPI;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)  
 DE (AdSS) (AMPSase).  
 GN PURA OR STM4366 OR STY4723 OR T4417.  
 OS *Salmonella typhimurium*, and  
 OS *Salmonella typhi*.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; *Salmonella*.  
 OX NCBI\_TaxID=602, 601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=*S. typhimurium*; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

Waterston R., Wilson R.K.;  
 "Complete genome sequence of *Salmonella enterica* serovar Typhimurium  
 LT2";  
 Nature 413:852-856 (2001).  
 [2]  
 SEQUENCE FROM N.A.  
 SPECIES=S.typhi; STRAIN=CT18;  
 MEDLINE=21534947; PubMed=11677608;  
 Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 Baker S., Basham D., Brooks K., Chillingworth T., Connor T.,  
 Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 Feltwell T., Hamlin N., Hogue A., Hien T.T., Holtroyd S., Jagels K.,  
 Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 Whitehead S., Barrall B.G.;  
 "Complete genome sequence of a multiple drug resistant *Salmonella*  
 enterica serovar Typhi CT18";  
 Nature 413:848-852 (2001).  
 [3]  
 SEQUENCE FROM N.A.  
 SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;  
 MEDLINE=22531367; PubMed=12644504;  
 Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
 Burland V., Kodyvanni V., Schwartz D.C., Blattner F.R.;  
 "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2  
 and CT18";  
 J. Bacteriol. 185:2330-2337 (2003).  
 CC -!- FUNCTION: Plays an important role in the de novo pathway of purine  
 nucleotide biosynthesis.  
 CC -!- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +  
 adenylosuccinate.  
 CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).  
 CC -!- PATHWAY: AMP biosynthesis; first committed step.  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the adenylosuccinate synthetase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AE008905; AAU23186.1; -;  
 CC EMBL; AL627283; CAD06843.1; -;  
 CC EMBL; AE016849; AA071866.1; -;  
 CC STyGene; SG??7??; purA.  
 CC HAMAP; MF\_00011; -, 1.  
 CC InterPro; IPR001114; Asucc\_synthase.  
 CC Pfam; PF00709; Adenylosucc\_synth; 1.  
 CC ProDom; PD001188; Asucc\_synthetase; 1.  
 CC TIGRFAMs; TIGR00184; purA; 1.  
 CC PROSITE; PS01266; ADENYLOSUCCIN SYN 1; 1.  
 CC PROSITE; PS00513; ADENYLOSUCCIN SYN 2; 1.  
 KW Purine biosynthesis; Ligase; GTP-binding; Metal-binding; Magnesium;  
 Complete proteome.  
 FT INIT MET 0 0 BY SIMILARITY.  
 FT NP BIND 12 18 GTP (POTENTIAL).  
 FT ACT\_SITE 140 140 BY SIMILARITY.  
 FT ACT\_SITE 147 147 BY SIMILARITY.  
 FT METAL 13 13 MAGNESIUM (BY SIMILARITY).  
 FT METAL 40 40 MAGNESIUM (VIA CARBONYL OXYGEN) (BY  
 SIMILARITY).  
 SQ SEQUENCE 431 AA; 47245 MW; EC7CA282DD51BF3B CRC64;  
 Query Match 7.7%; Score 84.5; DB 1; Length 431;  
 Best Local Similarity 24.3%; Pred. No.6.5;  
 Matches 51; Conservative 29; Mismatches 73; Indels 57; Gaps 11;  
 28 QDKGIAIPHDIDLGSRVVIQDY----DNQEQDRPTPS-----PAPSRPFSV-----L 72

Db 87 EDGIPVRERILLSEACPLLDYHVALDNAREKARGAKAIGTTGRCGIGPAYEDKVARGL 146  
 Qy 73 RANDVLWLSITAAEYDQSTYSGTGPVYVSDSTLVNVTGAQAVASRLDWTKTLDGRP 132  
 Db 147 RVGDL-----FDKETFAEKLKEYMEYHNFQLVNYK-----AEADVYQKVLDDTMA 192  
 Qy 133 LSTI-----QQVSKTFFVL--PLRGKLSFEAGTTKAGYPY--NVTNTASDQ 175  
 Db 193 VADLLTSMVVDVSDLLDQARQGRGDFVWFEGAQGTLLDIDHGT-----YPTVSSNTTAGG- 247  
 Qy 176 LLVENAAGHRVA-----ISTYTTSLGAGP 199  
 Db 248 VATGSGLGPRYVDVVLGILKAYSTRVCGAGP 277  
 RESULT 12  
 ENO\_BIFLO STANDARD; PRT; 432 AA.  
 AC OBG519;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-  
 glycerate hydro-lyase).  
 GN ENO OR BL1022.  
 OS Bifidobacterium longum.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
 OC Bifidobacteriaceae; Bifidobacterium.  
 OX NCBI\_TaxID=216816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCC 2705;  
 RX MEDLINE=22294977; PubMed=12381787;  
 RA Schell M.A., Karmirantzou M., Snell B., Vilanova D., Berger B.,  
 Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,  
 Pridmore R.D., Arigoni F.;  
 RT "The genome sequence of *Bifidobacterium longum* reflects its adaptation  
 to the human gastrointestinal tract";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427 (2002).  
 CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +  
 H(2)O.  
 CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing  
 the dimer (By similarity).  
 CC -!- PATHWAY: Glycolysis.  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the enolase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AE014725; AAN24829.1; -;  
 CC HAMAP; MF\_00318; -, 1.  
 CC InterPro; IPR000941; Enolase.  
 CC Pfam; PF00113; enolase; 1.  
 CC Pfam; PF03952; enolase N; 1.  
 CC PRINTS; PR00148; ENOLASE.  
 CC ProDom; PD000902; Enolase; 1.  
 CC TIGRFAMs; TIGR01060; eno; 1.  
 CC PROSITE; PS00164; ENOLASE; 1.  
 CC PROSITE; PS00164; ENOLASE; 1.  
 CC LYase; Glycolysis; Magnesium; Complete proteome.  
 FT ACT\_SITE 155 155 BY SIMILARITY.  
 FT METAL 242 242 MAGNESIUM (BY SIMILARITY).  
 FT METAL 285 285 MAGNESIUM (BY SIMILARITY).  
 FT METAL 312 312 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 432 AA; 46515 MW; B2F95935B262DE4 CRC64;

```
Query Match 7.7%; Score 84; DB 1; Length 432;
Best Local Similarity 19.5%; Pred. No. 7.2;
Matches 45; Conservative 25; Mismatches 69; Indels 92; Gaps 10;

QY 4 YSRPVWSANGPTVKLYTSVNAQDQKGIATPHDIDLGESRVVYQDYNQHQDRPTPSP 63
D 8 YARQLDSRGNTQVVLDTEDGAGLGL-VPSGASTGEAA----- 48
QY 64 APSRPFVLRANDVLMWLSLTAEDYQSTGSGTGPVYVSDSVTLVNVATGAQAVA----- 118
D 49 -----W-----ERDDGKSVYGGK-----VLNAVKAVNEVIAPKVGMDAAD 86
QY 119 RSLDWTQKVLDTGRPLSTIQYQSKTFVFLPRGK-----LSFWEAGTTAGVY-Y 166
D 87 QRALDDLMIELDG-----TPNKGKLGANAILGVSLAALYASAGLPLY 131
QY 167 NYNTTASDQLA-----VENAAGHR-----VAISTYTTSLGAG 198
D 132 RYIGGTNGHILPVPMNMINGAHADPATDIOEYMWISPYGFTYSEALRAG 182

RESULT 13
FXK1 MOUSE STANDARD; PRT; 617 AA.
AC P42128; O35939;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Forkhead box protein K1 (Myocyte nuclear factor) (MNF).
GN FOXK1 OR MNF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=94277065; PubMed=8007964;
RA Bassel-Duby R., Hernandez M.D., Yang Q., Rochelle J.M.,
RA Seldin M.F., Williams R.S.;
RT "Myocyte nuclear factor, a novel winged-helix transcription factor
RT under both developmental and neural regulation in striated
RT myocytes.";
RL Mol. Cell. Biol. 14:4596-4605 (1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=97415602; PubMed=9271401;
RA Yang Q., Bassel-Duby R., Williams R.S.;
RT "Transient expression of a winged-helix protein, MNF-beta, during
RT myogenesis.";
RL Mol. Cell. Biol. 17:5236-5243 (1997).
RN [3]
RP INTERACTION WITH SIN3B.
RC TISSUE=Heart;
RX MEDLINE=20088666; PubMed=10620510;
RA Yang Q., Kong Y., Rothermel B., Garry D.J., Bassel-Duby R.,
RA Williams R.S.;
RT "The winged-helix/forkhead protein myocyte nuclear factor beta (MNF-
RT beta) forms a co-repressor complex with mammalian Sin3B.";
RL Biochem. J. 345:335-343 (2000).
CC -!- FUNCTION: Transcriptional activator that binds to the upstream
CC enhancer region (CCAC box) of myoglobin gene. Has a role in
CC myogenic differentiation and in remodeling processes of adult
CC muscles that occur in response to physiological stimuli.
CC -!- SUBUNIT: Interacts with SIN3B to form a complex which represses
CC transcription.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P42128-1; Sequence=Displayed;
CC Name=2; Synonyms=Beta;
CC IsoId=P42128-2; Sequence=VSP_001545, VSP_001546;
CC -!- TISSUE SPECIFICITY: Expressed in tissues and cells in which the
```

```
CC myoglobin gene is transcriptionally active (cardiac and skeletal
CC myocytes, brain, kidney. . .).
CC -!- PTM: Phosphorylated.
CC -!- SIMILARITY: Contains 1 fork-head domain.
CC -!- SIMILARITY: Contains 1 FHA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; L26507; AAA37529.1; -.
CC EMBL; U95016; AAB69641.1; -.
CC PIR; A56051; A56051.
CC HSP; Q63245; 2HPH.
CC TRANSFAC; T04216; -.
CC MGD; MGI:1347488; Foxk1.
CC GO; GO:0003677; P:DNA binding; IDA.
CC GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IDA.
CC InterPro; IPR000253; FHA.
CC InterPro; IPR008984; SWAD_FHA.
CC InterPro; IPR001766; TF_Fork_head.
CC Pfam; PF00498; Fork head; 1.
CC PRINTS; PR00053; FORKHEAD.
CC ProDom; PD000425; TF_Fork_head; 1.
CC SMART; SM00339; FH; 1.
CC SMART; SM00240; FHA; 1.
CC PROSITE; PS00657; FORK HEAD 1; 1.
CC PROSITE; PS00658; FORK HEAD 2; 1.
CC PROSITE; PS00039; FORK HEAD 3; 1.
CC PROSITE; PS00006; FHA DOMAIN; 1.
CC DNA-binding; Transcription regulation; Activator; Nuclear protein;
CC Phosphorylation; Alternative splicing; Differentiation.
FT DOMAIN 27 33 POLY-ALA.
FT DOMAIN 107 159 FORK-HEAD.
FT DNA BIND 288 379 RSPASPTHPG -> SAPASHTSHA (in isoform 2).
FT VARSPLIC 400 409 /FTID=VSP_001545.
FT VARSPLIC 410 617 Missing (in isoform 2).
FT VARSPLIC 410 617 /FTID=VSP_001546.
SQ SEQUENCE 617 AA; 65839 MW; A1083B28C709FC4A CRC64;

Query Match 7.6%; Score 83; DB 1; Length 617;
Best Local Similarity 22.4%; Pred. No. 14;
Matches 45; Conservative 31; Mismatches 79; Indels 46; Gaps 9;

QY 30 KGIAIPHDIIDLGESRVVYQDYNQHQDRP-TPSPA-----PSRPFVLRANDVLMWLS 81
D 427 EGSPIDPHDPLGSLASVPEY--RYSQAFGSPVSAQPMVAVPRP-SNLVAKPVAVMP 483
QY 82 LTRAEYDQSTYSGTGPVYVSDSVTLVNVAT-----GAQAVASRLDWTQVT 127
D 484 ASIVTSQOPS-GHAIHVQQQAPTVMVRVVTTSANSANGYILASQSGTGTSHDTAGTAVL 542
QY 128 LDRPLSTIQYQ-SKTFVFLPRKLSFWEAGTTAGVY--YNTNTTASDQLLVENAGH 184
D 543 DLGNEAAGLEKPTIAFATIPAAASRVITQTVASQWAPGSPRTHSHHPTAG----- 591
QY 185 RVAISTYTTSLGAGPVISAV 205
D 592 -----YTSYRAAPLPVRAV 606

RESULT 14
BGLX ECOLI
ID BGLX_ECOLI STANDARD; PRT; 765 AA.
AC P33363;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
```

